

US science policy still in limbo

Last week has sharpened the difficulty in the United States of striking a balance between the needs of basic research and of the wider community.

THIS week's argument in Washington over the future of government support for science in the United States has a striking and curious feature. For practical purposes, US policy on research is in almost exactly the condition decreed by the British government a year ago. There will be national priorities for civil research spending, and the public research agencies will be accountable to the government for their diligence in the pursuit of those goals. The difference is that in Britain, the goals will be set by the government machine, while in the United States the goals will be set more publicly, with the US Congress having a decisive voice.

In neither country is it clear how the orthogonal considerations of excellence in research and of relevance to national goals will be reconciled. In the United States, there will be a particular difficulty in telling whether the Congress or the agencies of the administration will be responsible for setting national priorities. There is even a danger (as there is in Britain) that the goals set for the research enterprise will either be intrinsically unattainable or frankly populist. Senator Barbara Mikulski and Dr Neal Lane, the director of the National Science Foundation, may disagree on the exact balance of their influence (see page 401), but Mikulski's appropriations committee will be able to ensure that she has the last word.

It is unfortunate that this issue should have come to a head in the wake of the publication by the Office of Science and Technology Policy (OSTP) of what is meant as a framework in which the US research community will operate for the next several years. No doubt the community will be encouraged to know that President Bill Clinton and Vice-President Al Gore have their hearts in the right place, jointly supporting the objective that the United States should spend three per cent of its annual product on research. But researchers know as well as Clinton and Gore that there is no extra money about. The outlook is not bright.

So it would have been more prudent if OSTP had tackled some of the urgent questions that need attention. "Three per cent" or otherwise, basic research will be scrabbling for support in the United States for many years to come. OSTP would have done a valuable service if it had set out to define the scale below which untied research cannot safely shrink. That, in turn, would have provided a sheet-anchor for the research community.

The argument is simple. The United States needs people qualified to carry out research. The best and cheapest way to educate them is to engage them on research projects, prefer-

ably at academic institutions. Most of those thus educated will finish up in industry and the public service. Even in the United States, it may be prudent to provide graduate students with stipends so as to secure the attention of the brightest and the best. To have defined the irreducible support for basic research required to serve this educational function would have been a worthwhile national objective in its own right. Even Mikulski might have endorsed it.

The research community in the United States has also recently been outstandingly successful at uncovering truths about the world not previously suspected. That is an uncovenanted benefit, but one often achieved by too great a separation between research and teaching. Wise academic researchers appreciate already that the balance will have to be adjusted. OSTP might usefully have exploited that tendency to its own advantage and that of its dependent universities. It might also have put forward some convincing continuing role for the national laboratories, most of which have long outlived the original reasons for their existence.

By default, the prospect now is that the US research enterprise will be harnessed to implausible national goals. It is still within living memory that the National Science Foundation saddled itself in the 1960s with the programme called "Research Applied to National Needs", which came close to promising not only to predict earthquakes, but to prevent them (by injecting fluids into faults). The plan, now canvassed in the Congress, to make a research goal of the improvement of the physical infrastructure of the United States is, for different reasons, equally misplaced. What has the research community to contribute to the building of roads, bridges and houses that will not be done by the development departments of the great construction companies once there are funds to spend on construction contracts? □

The use of French

The French government's language bill is back in the melting pot and should remain there.

M. JACQUES Toubon, the French Minister of Culture, appears to have been as surprised as anybody that the constitutional council has now struck down his law on the French language (see *Nature* 370, 316; 1994), adopted only recently by the National Assembly and the Senate. But he should not have been surprised. Modern France is founded on a ringing



declaration of the centrality of personal liberty, much emulated elsewhere in the past two centuries. Would it not have been anomalous that a republican government that does not dare instruct its citizens how, for example, to dress on the beaches of Southern France would think of telling them how, and how not, to use their language?

None of that implies that Toubon's anxiety is misplaced. On the contrary, it is a proper part of a culture minister's brief to seek to safeguard the national language. But French deserves special regard, if not for the reasons usually given. French native-speakers say that their language is especially pleasing to the ear, but that is a prejudice of all native-speakers, as justifiable in Italy or Russia as in France. French also embodies a great volume of the world's outstanding literature, but who would claim that de Maupassant is more important than Dostoevsky or Dumas more important than Tolstoy? Translation is in any case a second best, while the deconstructionists would be quick to argue that translations of great works of literature lack the essential link between meaning and the language used to express it. It is the same with scholarship: those who wish to know what others have written have to learn the language.

The special case for French is that it is politically an important language. Many of those watching tragic television films from Rwanda in the past few weeks will have been struck that those speaking non-native languages there most often speak French (which they may have learned from the Belgians). In at least half of Africa, French is the first alternative language.

The difficulty with this view of the French or any language is that it evokes chauvinism. Any government's policy on its national language should be aimed at separating the cultural and social value of a vernacular that is easily and joyously used by voters in their mutual communication from the role of the same language as an instrument of foreign policy. To be fair to Toubon's law, it was mostly concerned with the use of French within France. Its faults were that it was unreasonably and (it now appears) unconstitutionally prescriptive. The British press has been making hay with the news that French citizens will not now be forbidden to welcome 'le weekend'; the question that nobody asked is whether such a requirement could have been enforced. The French government may be lucky that the constitutional council has come out against the law.

The dilemma remains, and indeed will be sharpened if the law is redrafted for another run through the Assembly. How to keep the language faithful to its roots and yet still widely used? That is an impossible trade-off. The Irish government has attempted that with Gaelic, and now has one of the purest and least used languages in the world. It would be a great misfortune for us all if French went that way. So the French government's best strategy would be to fight the intrusion of foreign phrases by ridiculing those that make no sense on the grounds that they are pretentious, but otherwise to let their language find its own level. English has become so widely used (and no worse a language) for just that reason, that there have been no constraints on how it is used by the great diversity of those who speak or write it. □

More bad news on AIDS

Japan is the site of the 10th International AIDS conference and the news seems more depressing than ever.

SCIENTISTS who keep themselves well informed about AIDS research know that progress in understanding such things as mechanisms of HIV replication have yet to produce a cure for the disease. Epidemiologists also know that HIV infection is about to reach major epidemic proportions in much of the world, particularly in Asia where heterosexual transmission is the primary mode of transmission.

In fact, from the incalculable amount of news print and air time that has been devoted to AIDS during the past decade, it is possible to conclude that there is no one alive who could be unaware of the dangers of unprotected sexual activity and the risk of contracting this lethal disease.

Alas, the data being reported from the 10th International AIDS conference in Yokohama suggest an astonishing lack of comprehension of the magnitude of the problem, let alone its biomedical solution. A report in the 11 August issue of *The New England Journal of Medicine* (331, 341–346 and 391–392) is equally discouraging. First, Japan. In what is said to be a show of solidarity and concern, the conference was attended by both the prime minister and the crown prince, but they had little to offer beyond a non-specific pledge to establish AIDS research and prevention programmes. This, coming in 1994, is too little, too late. According to the World Health Organization, the infection rate in Asian nations neighbouring Japan run to the hundreds of thousands, yet Japanese health officials record testing no more than 4,000 people at risk.

Equally disconcerting are data from a longitudinal study of AIDS transmission among heterosexual couples in which one partner was known to be HIV-positive. Reporting in *The New England Journal of Medicine* on behalf of the European Study Group, Isabelle de Vincenzi notes that all of the study participants were advised to use condoms during intercourse and that, among the 124 (or 48 per cent) of couples who followed that advice consistently, no seronegative partner contracted AIDS during roughly 24 months of follow-up. That's the good news. The bad news is that, despite knowledge about AIDS transmission, more than 50 per cent of the couples failed to use condoms consistently. Twelve of the seronegative individuals in that group became HIV-positive during the study.

The authors admit their study, one of the first to follow transmission patterns prospectively, is small. And this report is but a brief condensation of the data. Nevertheless, the message is clear. After all these years, the message about the danger of AIDS is not getting through. It is not just the medical scientists who have their work cut out for them. It is now painfully clear that the behavioural scientists had better put all of their talent into gaining a better understanding of human behaviour and how to change it before whole populations of relatively young men and women succumb to AIDS. □

Congress and administration differ over views of threats to science

Washington. Leaders of the US scientific community, eager for reassurance in difficult times and impressed in particular by the strong personal support of Vice President Al Gore, last week gave a warm welcome to the administration's science policy statement, *Science in the National Interest* (see *Nature* 370, 317; 1994).

But the most notable specific pledge in the document — a half-hearted promise to raise science spending to three per cent of gross domestic product (GDP) — was quickly shot down in flames in Congress, while sceptics complained that the paper does little to protect science from the growing propensity of Congress to direct scientific research.

According to White House officials, the document's assertion that "a reasonable long term goal for the total national research and development [R&D] investment (both civilian and military) might be about three per cent of GDP" was not meant to be a hard target. But it quickly became one.

At the launch of the document last Wednesday (3 August), for example, M. R. C. Greenwood of the Office of Science and Technology Policy (OSTP), chose to focus on the three per cent goal, while adding that achieving it was going to require "an unprecedented effort by government and the private sector".

But it did not take long for George Brown (Democrat, California), chairman of the House of Representatives Science, Space and Technology Committee, to remind

Greenwood just how unprecedented the effort would have to be. At a hearing the following day, Brown produced some back-of-an-envelope calculations suggesting that in order to achieve the target, the government's spending on civil R&D would have to grow by 80 per cent — to \$47.75 billion — by 1998.

"I concur with you that it's a challenge," John Gibbons, director of OSTP, told Brown rather lamely. "We're very interested in seeing how some of these figures work out."

The three per cent target was included in the document because OSTP badly needed some specifics to go with all the fine words.



Mikulski: seeking control of NSF?

But some remain unimpressed by a policy statement which, in the harsh words of one Senate staffer, "hasn't got any policy in it". The last such document, presented by President Jimmy Carter in 1979, brimmed with promises of more cash, especially for energy research. This time the OSTP says that the intent is different, namely to make a public case for science. Greenwood describes it as "a call to arms" for a scientific community that perceives itself as under threat.

Opinions differ on whether the threat is

real. There are two main funding agencies for basic research in the United States. Funding for the National Institutes of Health (NIH) is virtually guaranteed by congressional support for its work. So the smaller National Science Foundation (NSF) is the key battleground for science policy; and it is still not clear whether its independence is under threat from Congress.

NSF officials feel that the issue has receded since last autumn's attempt by a key Senate appropriations subcommittee to direct more of the agency's work towards strategic research. Since then, Barbara Mikulski (Democrat, Maryland), who chairs the subcommittee, has warmly praised the agency's director, Neal Lane, and backed large budget increases.

But Robert Park, of the American Physical Society, claims that the battle for the soul of NSF is continuing — and that Mikulski is winning. He points out that an NSF reauthorization bill, drafted chiefly by Mikulski and likely to be passed by the Senate, will require 60 per cent of the agency's funds to be devoted to "strategic goals".

These include favourites such as environmental research alongside less familiar entities, such as "civil infrastructure", which covers research intended to help repair crumbling roads and sewers and attracts the largest increases — 20 per cent a year, compounded over five years — in the bill.

The bill is unlikely to pass into law, as it is significantly different from the version under consideration by the House. But the approval of Mikulski's version by the Senate alone will lend authority to her committee's future budget decisions.

The change means that NSF money will be authorized in two ways, by mission and by discipline. NSF grant applications are already being categorized in this manner. But Park thinks the change is dangerous, because although Congress has tended to leave disciplines alone, it will feel free to tamper with missions. "It gives Mikulski control over NSF," he says.

Lane himself points to language in the bill which gives him, as director, power to revise the strategic goals. He says that the bill gives the NSF sufficient flexibility to readjust the areas concerned, and that such areas "are precisely the ones which we, at NSF, identified".

The impact of both Mikulski's reauthorization bill and the White House policy document will be measured by the principles they establish, not the dollars they specify. And, at present, it seems to be Mikulski who is making the running.

Colin MacIlwain

Fast reactors 'could rise from ashes'

Paris. Superphénix, the world's largest prototype fast breeder reactor, was kicked into operation last week at Creys-Malville near Lyons for the first time in four years, just days after the French government had approved its conversion from a power station into a research reactor.

The decision to restart the 1,200-MW reactor appears primarily to have been a political solution to the need to satisfy NERSA, the French, German and Italian consortium that owns the reactor. NERSA has paid FF27.7 billion (US\$5 billion) to build the reactor, which has operated for a total of less than six months; it was shut down in 1990 after repeated leaks in its sodium cooling circuits.

Half of the programme's FF100 million annual budget will be spent converting the fast neutron reactor from a breeder to a consumer of plutonium. This will be done

by equipping it with cores lacking a uranium blanket, within the CAPRA (Consommation Accrue de Plutonium dans les Rapides) programme started by the French Atomic Energy Commission last autumn.

A third of the funding will go to studies on incinerating minor actinides, such as neptunium, and one-fifth to safety studies.

Adrien Mergui, president of the directors of NERSA, claims that the restart of Superphénix will help to restore some of the lost momentum in the development of fast breeders. He predicts a resurgence of interest in fast breeders, given the inevitable — if more distant than previously been expected — shortage of uranium resources. The planned research at Superphénix, he says, will provide the groundwork for the European Fast Reactor programme, abandoned last year, to rise from the ashes.

Declan Butler

See also page 404.

Search launched for lowest 'addictive' nicotine level

Washington. Last week's conclusion by an advisory panel of the US Food and Drug Administration (FDA) that nicotine is addictive has shifted the scientific focus of debate over the regulation of tobacco to the question of what dose sustains the addiction.

The conclusions of the FDA's drug abuse advisory committee may seem a statement of the obvious, although they are still disputed by scientists for the Tobacco Institute, which represents the views of the tobacco industry, who argue that nicotine does not intoxicate — and that intoxication is a defining symptom of addiction.

The FDA panel rejected that line of reasoning. At a meeting in Washington, DC, it agreed that nicotine's addictiveness is "substantially due to the pharmacological effects of nicotine on the human body". But it was unable to agree on the level of nicotine that is addictive.

The findings are critical to the FDA's current attempts, led by its commissioner, David Kessler, to assume jurisdiction over all tobacco products.

One of the main requirements for seeking such authority under existing legislation is that the substance changes the structure and function of the body; hence the FDA's focus on nicotine, even though other constituents of tobacco are more harmful.

But as the law now stands, if the agency were to assume jurisdiction it would have to ban tobacco products until the industry could prove their safety and efficacy. Such action would provoke a fierce political and social outcry. Legislation is currently before Congress that would allow the FDA to give a more moderate response — for example, to require new labels, spelling out the chemical content and addictive nature of cigarettes.

The phrasing on such labels may, in turn, require the FDA to decide whether — and if so, at what level — there is any minimum dose leading to addiction. In the longer term, more drastic action is likely to follow. One congressional source anticipates that "within the next year or two Kessler will ban cigarettes with more than a particular level of nicotine and tar".

Such a ban would not cover all cigarettes containing addictive levels of nicotine. But it could lead to congressional approval of a variation of a strategy proposed by Neal Benowitz, chief of clinical pharmacology and experimental therapeutics at the University of California, San Francisco, and Jack Henningfield, chief of the clinical pharmacology branch of the National Institute of Drug Abuse (NIDA).

At last week's meeting, Benowitz and Henningfield suggested that the nicotine

content of cigarettes should be reduced over 10 to 15 years to a level that does not maintain addiction. They argue that approach would both minimize the burden imposed by the 45 million Americans currently addicted to nicotine and reduce the number of new smokers.

But if this approach is adopted, it will require a consensus on the maximum level

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Details of 'addiction' levels may soon be required on cigarette packaging.

of nicotine accepted to be non-addictive. Benowitz and Henningfield suggest 5 mg a day, on the basis of an analysis by Saul Schiffman, a psychologist at the University of Pittsburgh, of the roughly 10 per cent of smokers who are not addicted — those who can smoke one or two cigarettes every few days and stop without withdrawal symptoms.

A randomized double-blind trial of nicotine patches by David Sachs, director of the Palo Alto Center for Pulmonary Disease Prevention in California, suggests that the 5 mg figure may be right.

Sachs, Benowitz and Henningfield all accept that more research is needed. Their studies are based on existing smokers, and their data may therefore not accurately reflect the dose needed to create an addiction.

Partly as a result of the FDA's recent action the National Cancer Institute (NCI) and the NIDA are likely to put out joint requests for research aimed at establishing the minimum dose of nicotine that is addictive, according to Thomas Glynn, associate director for cancer control science at the NCI.

Benowitz and Henningfield are themselves each seeking a physiological signal, such as a brain-wave pattern, that may indicate susceptibility to nicotine addiction. If such an indicator is found, pure nicotine could be administered to volunteers through a patch, and stopping long before an addiction develops. This would provide a measure of an average level of sensitivity to nicotine, from which regulators could establish a non-addictive dose for each cigarette.

Helen Gavaghan

Pay-by-results leads to funding shift in Australian universities

Sydney. A new performance-based system of allocating research funds to Australian universities, whose results were announced last week, has shifted funds away from the established institutions to a number of newly created universities.

As the new system rewards links with industry, it has also shifted millions of dollars away from institutions in relatively unindustrialized parts of the country, such as the University of Adelaide (in South Australia) and the University of Western Australia, towards those in the more heavily industrialized east coast.

The biggest loser, Adelaide University, will see its annual grant cut by A\$3.8 million (US\$2.8 million) by 1996. In anticipation of the cut, it has already decided to shed 90 staff, including 15 researchers, through voluntary redundancies.

The new funding system involves about 6 per cent (A\$248 million), of the general government grants provided to all Australian universities. The federal government has decided to allocate this sum as a "research quantum", allocated according to a 'composite index' that takes into account a range of 'input' and 'output' factors.

Inputs include the institution's success in winning research grants from funding bodies and from industry. Outputs include the number of publications and postgraduate completions.

One effect of this process has been to give more money to newer universities, many of which have been successful in attracting industry funding for research.

The University of Western Sydney, for example, founded in 1989, will see its share of the research quantum increase from a meagre A\$300,000 to around A\$1.4 million a year by 1996. The Queensland University of Technology (formerly the Queensland Institute of Technology), will see its share of the research quantum rise from A\$1.3 million to A\$2.7 million a year in 1996.

Professor Brian Wilson, vice-chancellor of the University of Queensland, which will receive an additional A\$2.7 million a year by 1996, claims that the university's success is based on its record in attracting both industry funds and government research grants in agriculture and mining.

In contrast, Mary O'Kane, deputy vice-chancellor of research at the University of Adelaide, says that the university had suffered in the fight for industry funds because most Australian companies are based on the east coast. She said that the eventual cut of A\$3.8 million a year in the university's operating budget will be difficult to absorb, and will force it to look elsewhere for funds, particularly to Asia.

Mark Lawson

Fusion meets the realities of global politics

Paris. Scientific and engineering goals must be tailored to political needs if international collaborations on 'big science' projects are to work. That seems to be one message emerging as the dust settles on a recent management crisis at the International Thermonuclear Experimental Reactor (ITER) project, which last month cost the director, Paul-Henri Rebut, his job.

ITER, formally launched in 1988, was planned from the outset as a collaboration between Europe, the United States, Russia and Japan, aimed at demonstrating the feasibility of controlled nuclear fusion (see *Nature* 358, 269; 1992). It is intended to produce a design for a huge tokamak reactor capable of sustaining fusion in a magnetically confined deuterium-tritium plasma for 1,000 seconds, while demonstrating that the process is safe.

A decision whether to extend the ITER agreement to include construction will be made before 1998 — the current engineering design phase commits the partners only to producing sufficiently detailed designs by then to allow construction "either alone or together". If the reactor is eventually built, it may be followed by a commercial prototype around 2025.

So far, the \$5.6-billion ITER collaboration has been widely considered a model of how to manage a 'big science' project. But after Rebut was dismissed by the ITER council — which is made up of representatives of the four participating blocs — he complained that the "diverse priorities of the national programmes will not result in a coherent engineering design".

In particular, Rebut argued that the ITER agreement needed to be restructured to give his central team the "responsibility and authority" he felt was needed to do the job properly. One observer, for example, says that Rebut, who had little power to allocate funds and research, feared that design by a quadripartite committee would result in a "dog's breakfast".

Rebut said that during his previous experience as the widely-respected head of the Joint European Torus (JET), at Culham in the United Kingdom, the "project made the key decisions, and the national programmes played a supportive role". He has argued that: "The same must also be true for ITER."

But politics rather than efficiency seems to have determined the division of ITER research among 'co-centres' in Garching in Germany, Naka in Japan, and San Diego in the United States. Some claim that this distribution was the price of securing collaboration, accepting that, as ITER is a paper design and not an engineering project, a single site is not strictly essential at present.

Fusion supporters are divided over the legitimacy of Rebut's complaints. JET's internal newsletter *JET News* implicitly sup-

ported his call for a restructuring, and speculated that he might have been fired for not "suffering fools gladly". It described Rebut as a "brilliant engineer/physicist and a decisive leader", who would "undoubtedly be a hard act to follow".

Others argue, however, that Rebut failed to appreciate the differences between managing JET and coordinating ITER. One official from the European Union (EU) fusion programme says, for example, that Rebut wanted to design a much more ambitious reactor than that originally agreed to by the four partners.

The official claims that Rebut wanted ITER to demonstrate that fusion was both technologically and economically feasible. He says that Rebut's design would have pushed ITER well over its agreed budget, and required technology that does not yet exist — for example, reactor blankets that could breed further tritium to reintroduce as fuel. (Rebut is currently on holiday, and could not be contacted for comment.)

A management review carried out by the ITER council concluded in May that Rebut's management of the programme was too hands-on. "Rebut ran ITER as a one-man show", thus denying researchers in the four blocs the participation in ITER that they had paid for says one EU official. "Rebut thought ITER was a blank cheque for his own dreams."

Robert Aymar, Rebut's successor, says that Rebut failed to appreciate that he did not, and could not, have the same degree of control over ITER as he had over JET. The latter is four-fifths funded by Euratom (the

shaping the various contributions of the projects and industries of the participating countries into a coherent "single project".

One difficulty in persuading groups to cooperate within ITER, says Aymar, is that it is a paper design project. This means that laboratories often prefer to further their own projects, as they can get more immediate scientific results from their own smaller but operational equipment.

Aymar says his main aim as director will be to ensure that ITER is actually built. One EU official says that if it is not, the partners — who together spend some \$1.5 billion annually on fusion research — "will say fusion is dead". Indeed, a bill passed in the US Senate already proposes that if ITER fails, the United States should reduce fusion funding from \$344 million to \$50 million.

The four blocs are collaborating on ITER only because none of them — with the possible exception of Japan — is sufficiently convinced of the technological and economic feasibility of fusion to provide the heavy development costs on its own. Although some EU officials say that Europe says it would go it alone if ITER failed, others are sceptical.

To ensure ITER's success, Aymar is prepared to be pragmatic. He says, for example, that he will settle for conservative choices in the design of ITER to keep the projects within its current budget estimates, "even if these are not the best for fusion in the long term".

Some sponsoring governments anticipate that, when a site for the reactor is finally chosen, the national blocs not selected may lose interest and fall back on their national programmes.

To avoid this happening, Aymar says the decision on whether to construct ITER, how much each country will pay, and where the reactor will be sited, will be negotiated within a single package, sometime between 1996 and 1998.

One side-effect of the management crisis at ITER has been to reopen public debate on the wisdom of fusion research. But whereas the publicity given to reports of cold fusion in 1989 unexpectedly increased awareness and support for hot fusion, the latest publicity has reawakened the critics.

The European Parliament, for example, which has long been sceptical of claims that fusion can offer limitless cheap, clean energy, recently nominated a critic of fusion research, Detlev Samland (Social Democrat, Germany), to head its powerful budget committee.

With pressure growing on the EU to spend more on research areas of immediate economic relevance, the ECU200 million spent annually on fusion research may become an increasingly attractive target for cuts.

Declan Butler



European Atomic Energy Community) whereas ITER has to accommodate the needs of the programmes of four powerful blocs.

Aymar agrees that this quadripartite structure "cannot be as effective as a small team". But he insists that such structures are "inevitable" in international technological collaborations.

The job of the ITER director, Aymar says, is to make the best of this situation by

US reduces options on disposal of plutonium

Washington. The US Congress has killed the experimental Integral Fast Reactor (IFR) at the Argonne West National Laboratory, Idaho, reducing the likelihood that nuclear physics will play a role in disposing of the United States' plutonium stockpile.

Last week, a conference of House and Senate members agreed to eliminate \$100 million annual funding for the IFR. This will end work on actinide recycling, a technique aimed at reprocessing plutonium into a radioactive (and militarily useless) compound of heavy metals.

The technique was the most advanced of a set of ideas that had been proposed by nuclear physicists to deal with plutonium. But a recent report from the National Academy of Sciences concluded that its application was too distant to help deal with the immediate plutonium problem (see *Nature* 367, 307; 1994).

Furthermore, efforts by scientists at the

Los Alamos National Laboratory in New Mexico to promote a rival approach, using particle accelerators to destroy plutonium and long-living radioactive isotopes, appear unlikely to attract serious interest from the Department of Energy.

At the first international conference on so-called Accelerator Driven Transmutation Technology (ADTT), held in Las Vegas two weeks ago, Los Alamos announced a \$3 million research collaboration with Russia on the transmutation of plutonium.

But no high-level Department of Energy officials attended the meeting. And industry representatives said privately that the laboratory had little chance of securing the \$500 million it wants for an ADTT demonstrator. Around \$25 million — mostly from the nuclear weapons programme — has already been spent there on studies of the technology.

Charles Bowman, head of the ADTT office at Los Alamos, concedes that trying

to sell the technology to the National Academy study team, as well as to another group of scientists from the Department of Defense's (DoD)'s Jason Program who also reported unfavourably on it in January, has been "a challenge". "They say the technology isn't near enough to deployment to deal with present, urgent problems: we don't agree," says Bowman.

The Las Vegas meeting was an attempt to build public, as well as scientific, support for transmutation technology, which gained widespread publicity late last year with the endorsement of Carlo Rubbia, former director of CERN, the European Laboratory for Particle Physics (see *Nature* 366, 392; 1994).

"This is a programme whose time has come," says John Brown, a division director at Los Alamos. But, he adds, "these are very difficult times for new technological ideas. The public is sceptical about all technologies, not just this one."

ADTT would use a proton accelerator to bombard a target, creating enough neutrons to sustain fission in a subcritical nuclear reactor. Bowman says that the extra neutrons would enable the machine to keep on destroying undesirable fission products which would poison conventional nuclear reactors.

All of this would take place in a molten salt coolant-fuel which would be chemically recycled in a closed loop, ploughing plutonium and long half-life isotopes back in until they transmute, and therefore generating no long-lived nuclear waste.

This process would generate vast amounts of radiation from the short half-life isotopes produced. But Bowman thinks such radiation could be contained in the closed cycle, where separation would take place under remote control.

The physics of the process dictates that one transmuting power station would be needed to deal with the plutonium produced online by four existing nuclear power stations. A vast construction programme would therefore be needed to deal with new plutonium — as well as the backlog.

Despite these daunting obstacles, Carlo Rubbia describes the process as "a sensible proposal" which he said that scientists should be looking in to.

But, as Bowman concedes, the easiest thing to do with ADTT is to use it to make tritium for use in atom bombs. Now that the reactors which used to do this have been closed for safety reasons, the DoD needs such a facility to help maintain its 20,000 nuclear weapons.

The DoD is pressing for a speedy decision. Pete Domenici, the Republican Senator for New Mexico, also known as 'the Senator for Los Alamos,' wants it to be an accelerator-based facility, built at the weapons laboratory. **Colin Macilwain**

Allain freed — to face new charges

Paris. Jean-Pierre Allain, the haematologist sentenced to four years' imprisonment (with two suspended) in 1992 in the French contaminated blood affair, was indicted last week on poisoning charges, less than 24 hours after a parole board had announced that he would be released the following day. He was eventually released on Monday evening.

A week earlier, the Paris Bar Association, prompted by the indictment — also on poisoning — of Michel Garretta, former director-general of the French National Transfusion Centre (CNTS), had warned the French legal system against letting public opinion override the law in its prosecution of individuals alleged to have supplied haemophiliacs with factor VIII contaminated with HIV in the mid-1980s (see *Nature* 370, 319; 1994).

The timing of Allain's indictment has provoked similar widespread criticism. *Le Monde* described it as "the cruel image of a cat squashing a mouse with a blow of its claw after letting him think he was free", adding that the French legal system does not usually behave like this.

Despite the new charges pending, a court agreed on Monday to free Allain on parole. Among those waiting outside the courtroom was Robin Carrell, head of the department of haematology at the University of Cambridge — where Allain holds a faculty appointment — who has led an international campaign to prove Allain's innocence.

The court has said that one condition for Allain's freedom is that he should stay in France, and he will therefore be unable to return to Cambridge, where his

post is being kept open.

Many scientists worldwide have backed Carrell's campaign, and have sent letters of support to Allain. More than 30 Nobel prizewinners, for example, have written individually to President François Mitterrand to plead Allain's case.

Allain's wife, Helen Lee, who worked with him at CNTS, describes the new charges as a "nightmare". She is convinced her husband has been made a scapegoat to appease public opinion, and that the justice system has deliberately ignored evidence that he is innocent.

Indeed, Jean-Claude Gluckmann, professor of immunology at the Pitié-Salpêtrière Hospital in Paris, submitted written evidence to Allain's appeal in 1993 stating that either Allain was innocent, or many others — including himself — were also guilty. His memorandum was ignored, he says.

Lee accuses many other leading French AIDS experts and physicians of "hypocrisy" and "cowardice". She is particularly distressed that whereas many now declare they knew of the risks of HIV contamination in 1985, but that their warnings were ignored, quotations from the same individuals in newspapers from 1985 show that they, too, had failed to anticipate the danger. Declan Butler

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Allain: now accused
of 'poisoning'.

Frank Spooner

Energy moves 'still insufficient to meet carbon goals'

Tokyo. Japan is unlikely to meet an international commitment to reduce total emissions of carbon dioxide to the 1990 level by 2000, according to a report released last week by the Environment Agency. But it may be able to keep its already low per capita emissions to the 1990 level.

In 1990, the Japanese government set two different targets of keeping total emissions and per capita emissions of carbon dioxide to the 1990 level (see *Nature* 347, 700; 1990). This was a compromise between the strongly differing views of the Ministry of International Trade and Industry (MITI) and the Environment Agency.

The per capita goal, preferred by MITI, is less ambitious and allows for a slight increase in total emissions because of the increase in population. The second, more stringent goal was set in preparation for the 'Earth Summit' in Brazil in 1992, at which developed nations agreed to produce plans for reducing total CO₂ emissions to the 1990 level by 2000.

The Environment Agency's report is based on long-term projections of energy supply and demand made recently by MITI's Agency of Natural Resources and Energy. Providing that it is approved by the cabinet, the report will be submitted next month to the convention secretariat, based in Geneva.

Since 1990, Japan's total carbon dioxide emissions have increased one or two per cent a year. If this trend continues, according to estimates by the Environment Agency, its total emissions will increase by nearly 14 per cent, to 364 million tonnes, by 2000.

According to the Agency of Natural Resources and Energy, a thousandfold increase in the use of solar power, as well as extensive energy conservation in the home, industry and transport, would reduce CO₂ emissions by about 30 million tonnes over this period.

The new Environment Agency report says that, if these goals are achieved, and other measures to reduce energy consumption are implemented, the results would be an increase of only 3.1 per cent on 1990 — and a reduction in per capita emissions to 2.6 tonnes per year, the same as 1990.

Although this level is already one of the lowest in the developed world, some European countries, such as Germany and Denmark, are already talking of committing themselves to even lower targets. The Environment Agency report says Japan must therefore make "added efforts" to stabilize total emissions at the 1990 level. But neither MITI nor Japanese industry, both of which are worried about the impact of restrictions on economic growth, seem likely to support this call.

David Swinbanks

Japan shifts space emphasis towards unmanned missions

Tokyo. Japan is scaling down its plans for long-term space development to focus particular attention on the Moon and Earth, with a switch of emphasis from manned missions to unmanned observation and exploration. But the plans remain ambitious.

Japan is also planning to propose next month a project to establish an international global Earth observation system with about a hundred satellites. Japan's own proposed contribution is likely to be 20 to 30 Earth observation satellites launched by 2010.

At the end of last month, the 'special committee on long-term vision', an advisory committee to the Space Activities Commission, submitted its recommendations for the period up to and beyond the year 2020. These will be used to revise government policy before the end of this year.

The report differs from one published seven years ago (see *Nature* 327, 645; 1987) chiefly in the absence of recommendations for either a manned space station or manned space shuttle to be built by Japan.

The increased emphasis on unmanned missions is in line with the views of the Institute of Space and Astronautical Science (ISAS), the smaller of Japan's two space agencies, which has achieved international recognition for its relatively modest scientific missions.

The report calls for systematic exploration of the Moon, using satellites and remote-controlled vehicles, with the aim of establishing an unmanned astronomical observatory between 2010 and 2020. This contains elements of recommendations for a Moon base published in June by the Lunar and Planetary Society (see *Nature* 369, 435; 1994), but with the key difference that the observatory will not be manned.

The proposals for lunar exploration over-

lap with plans of the European Space Agency. The report calls for international cooperation in establishing the lunar observatory, and Japanese space scientists say they hope to cooperate with Europe in lunar exploration, if the Japanese and European proposals are approved by their respective governments.

Smaller missions to the Moon and nearby planets, such as Mars and Venus, will be launched with the solid-fuel M-V rocket currently under development by ISAS. Larger missions, and missions to more distant planets such as Jupiter, will be handled by the new liquid-fuel H-II rocket of the National Space Development Agency, launched successfully for the first time earlier this year.

The report calls for the development of a larger H-II rocket, with double the payload capacity of the present version. This would be able to put a satellite weighing four tonnes into geostationary orbit, and would also be able to launch the 20-tonne unmanned space shuttle HOPE, expected to be built by the end of this decade.

The report's proposals will, if adopted, require an investment of ¥7,000 billion (\$70 billion) over 15 years from 1995, with a consequent increase of nearly 10 per cent a year in Japan's current space budget of about ¥200 billion.

This is an ambitious target, given the severe fiscal restraint policies of the Ministry of Finance, particularly during the current recession. But the space budget has recently been increasing by between eight and nine per cent per year, largely as a result of Japan's contribution to the US Space Station. The new budget requirement is therefore well within the realms of possibility.

David Swinbanks

Olympic science gets \$1/4 million prize

London. The International Olympic Committee (IOC) has announced the launch of a new prize intended to recognize "the evolution of scientific research related to human development". The prize, sponsored by the pharmaceutical company Parke-Davis, will include a medal and cash sum of \$250,000.

The IOC Olympic Prize will be awarded every two years for research in the biological, medical, physical or psychological sciences. The first winner (or winners) will be announced at the 26th Olympic Games, which take place in Atlanta, Georgia, in 1996.

Meanwhile, French mathematicians, worried at signs that their traditional

impact on world science may have been decreasing in recent years, are celebrating the award of Fields Prizes — the discipline's equivalent of the Nobel Prizes — to two of their colleagues, Pierre-Louis Lions and Jean-Christophe Yoccoz.

The third Fields winner, Jean Bourgain, although a native of Belgium, has lived in France for many years, and combines positions at the Institut des Hautes Etudes Scientifiques outside Paris and the Institute for Advanced Study at Princeton University in New Jersey. The fourth winner is the Russian-born Efim Zelmanov, now at the University of Chicago. □

'Cheap option' UK science courses under fire

London. Britain's Royal Society has added its voice to that of science teaching bodies in criticizing the government's curriculum advisers for not encouraging some 14–16-year-olds to study more science than the minimum stipulated in the national curriculum.

The criticisms form part of the society's comments on proposals drawn up by the government-funded School Curriculum and Assessment Authority (SCAA) for a broad-ranging reform of the curriculum.

Science is one of seven subjects that all pupils must study up to age 16 under the national curriculum. Single-award science, which is equivalent to a single pass at GCSE (examinations taken at the age of 16), is intended to cover the three principal sciences in the statutory minimum of 10 per cent of curriculum time, and is recommended by SCAA for a 'minority' of pupils with a good reason for focusing on other subjects.

But the Royal Society, the Association for Science Education and the National Union of Teachers all argue that this is a second-best option to the so-called 'double-award course', taught in 20 per cent of the curriculum, and that it seriously limits opportunities for further study of science.

The scientific bodies argue that the demands on the course mean that the coverage of topics that have not been dropped completely is "superficial".

The result, they say, is to leave such pupils ill-equipped to deal with issues "in a society where science and technology are part of everyday life". And it closes off

further study of science after the age of 16, conflicting with the government's goal of a balanced curriculum.

The Association for Science Education (ASE) had already written to Sir Ron Dearing, chairman of SCAA, earlier this year asking for clarification of the government's intention after an interim report led to confusion over how much science was needed for the average child.

ASE officials claim that some head teachers had interpreted SCAA's recommendations as meaning that a majority of pupils should take the single-award option, and had seen this as a way of easing pressure on tight budgets and timetables.

Dearing reaffirmed in the proposals recently published for comment that SCAA expects "the large majority" of pupils to study double-award science. But the ASE says it remains "very concerned" that he did not recommend that all pupils between the ages of 14 to 16 should spend 20 per cent of curriculum time on science.

SCAA's review of the national curriculum, which was introduced in 1988, was commissioned in 1993 after the government received widespread criticism from teachers and academics. Its conclusions will be presented to Gillian Shephard, the newly appointed Secretary of State for Education, in the autumn, once the responses to a consultation document sent out in May have been analysed.

Both the National Union of Teachers and the ASE also say they are concerned that

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Concern is growing that some school science courses are too superficial.

general availability of single-award science will reverse the recent increase in the number of girls taking science after the age of 14, which they argue has been achieved through the national curriculum and balanced science courses.

Their concern is based on statistics showing that girls are in a slight majority of the 12 per cent of pupils taking single-science courses, and in a minority of the 82 per cent taking the double-science option. (The remaining 6 per cent of pupils still take the traditional "single subject" courses at GCSE.)

Maggie Verrall

Zeneca cuts seeds but enters genome technologies

London. The decision two years ago by the UK company Imperial Chemical Industries (ICI), in the wake of a failed hostile takeover bid, to split off its science-driven activities into a new company, Zeneca Group plc, seems to be paying off (see *Nature* **362**, 780; 1992).

Last week, Zeneca announced pre-tax profits for the first six months of this year, excluding exceptional items, of 31 per cent, considerably higher than the market had been expecting.

The company's shares were expected to receive a further boost this week with the announcement that Zeneca's subsidiary Cellmark Diagnostics is launching a long-awaited screening kit for cystic fibrosis.

The news is not all good. Zeneca has also announced that, as a result of the continuing 'economic disorder' in the countries of eastern Europe it is withdrawing from its hybrid seeds business in the region, a move that will involve writing off £100 million (US\$154 million) — and could lead to the loss of up to 500 jobs in Britain and elsewhere.

It has also encountered difficulties in

increasing sales in the United States of the anti-breast cancer drug tamoxifen, sold under the brand name Nolvadex, following controversy over the conduct of trials of the preventive use of the drug in the United States (see *Nature* **368**, 679; 1994).

Nevertheless analysts appear to be endorsing Zeneca's claims to have a healthy "product pipeline", with clinical trials well advanced on, for example, new drugs for the treatment of breast and colo-rectal cancer.

One of the factors behind the decision to split ICI (the bulk chemicals businesses are now concentrated in ICI plc) was to enable Zeneca to focus more sharply on developing the applications of advanced science-based techniques in pharmaceuticals and agrochemicals.

Peter Doyle, formerly research director of ICI and now an executive director of Zeneca, cites the company's recent announcement of a new variety of banana, genetically engineered to reduce its susceptibility to rot, as an indicator of progress in this direction.

Zeneca has recently come under some

criticism for its apparent tardiness — at least in comparison to rivals such as Glaxo and Roche — in entering the field of human 'genomics', in particular the gene-based diagnosis and treatment of diseases.

Tom McKillop, chief executive officer of pharmaceuticals, disputes these criticisms: Zeneca's recent decision to establish a human genetics group as well as the launch of Cellmark's cystic fibrosis kit underline its commitment to this field, he says.

The company is also believed to be preparing to announce a strategic alliance with an outside research group in human genomics. Such an alliance was announced two weeks ago between Glaxo and the San Diego 'gene-hunting' company Sequana.

For the time being, however, Zeneca's most public excursion into the domain of genome technologies has been the selection of Cellmark's US subsidiary by the Los Angeles police department to conduct the DNA profiling tests on evidence in the trial of O. J. Simpson, the former football star charged with murdering his former wife and her male friend. □

Astronomers and squirrels in new clash over telescope

Washington. A long-running feud between US astronomers and conservationists has erupted once more in the mountains of Arizona, delaying the construction of what will eventually be one of the world's most powerful ground-based telescopes.

The Tucson district court ruled at the end of last month that site preparation for the Large Binocular Telescope (LBT) on Mount Graham should be halted pending new environmental impact studies, following a decision to shift its site by 400 yards from that originally planned.

The University of Arizona — which is building the \$60-million telescope together with the Research Corporation of Tucson and an Italian consortium led by the Arcetri Astrophysical Observatory in Florence — plans to appeal against the ruling.

But this latest twist in a tortuous 14-year saga has increased concern among astronomers that environmental pressure to protect the wilderness may eventually make it impossible to build new telescopes on any desirable site in the United States.

Mount Graham was originally chosen as the site of an international observatory because, unusually for a suitably remote mountain in the dry American southwest, it already had some tourist development and good road access.

But, in the latest of ten court battles, Alfredo C. Marquez, a district court judge, ruled on 28 July that, by allowing the university to shift the site of the new telescope, the US Forest Service and Fish and Wildlife Service had breached a 1988 agreement allowing the observatory to be built.

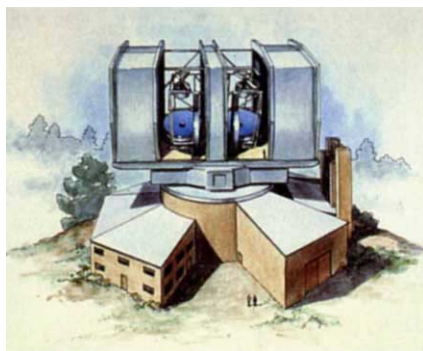
A coalition of local and national environmental groups brought the case to court when the university cleared 250 trees from the new site last December. They argue that the observatory already damages a small and retreating ecosystem at the top of Mount Graham, threatens the future of a now-famous subspecies, the Mount Graham Red Squirrel, and offends San Carlos Apaches, who consider the mountain to be sacred.

The LBT would be the third and largest telescope at the site. Sitting on top of a 100-foot tower on one acre of land in the middle of nowhere, this pristine machine does not appear to be the gravest threat facing the environment of the continental United States. But try explaining that to the red squirrel, says Roger Featherstone, a long-standing opponent of the project, who now works with the Endangered Species Coalition in Washington, DC.

Featherstone says that the top of Mount Graham consists of 600 unique acres of mixed spruce and fir, and describes it as "a little piece of the north Canadian ecosystem"

tem" in Arizona. "We're talking about a very tiny ecosystem," he says. "The third telescope would be the largest and most destructive."

The squirrel is one of 25 red subspecies found in North America. When it was listed under the Endangered Species Act in 1987, it was thought to survive only in the spruce and fir environment at the mountain top, and



The Large Binocular Telescope being planned by the University of Arizona.

the observatory had to obtain an exemption from Congress to build on the land.

But the University of Arizona, which employs five full-time biologists to watch the squirrels and pays \$50,000 a year to the Forest Service towards research on their habitat, now thinks they are also thriving further down the mountain.

Peter Strittmatter, director of the Steward Observatory, says he has been told that the new site for the LBT contains even fewer squirrels than the original one. But opponents say that the main point is that the university broke the law by shifting the site.

"At 4.30 in the morning, in the middle of the winter, they clear-cut those trees," says Bob Witzeman, a retired Phoenix physician and long-standing opponent of the project, who claims that the act was illegal. "We started at the crack of dawn as we always do," retorts Strittmatter. "The Forest Service gave us the authority to do it."

The argument will continue in court. The university last week filed an emergency motion asking the appeal court to lift the ban, hoping to do some work this year before snow falls.

The controversy has already cost the Mount Graham Observatory the support of its original sponsor, the Smithsonian Astrophysical Observatory, and of every North American university courted as a partner. But Strittmatter thinks the observatory would have faced similar opposition anywhere in the United States. "The irony is that observatory sites are always so benign, almost like wildlife refuges," he says.

Colin Macilwain

Congress directs NASA to plan an asteroid hunt

Washington. In the aftermath of what some are describing as "comet fever", the US Congress has directed the National Aeronautics and Space Administration (NASA) to develop a plan by next February for a long-term programme to identify comets and asteroids that may pose a threat to Earth.

In an amendment to next year's authorization bill for the space agency, the House of Representatives Science, Space and Technology Committee has instructed it to produce a strategy — including funding requirements up to 2000 — for a programme capable of identifying all objects larger than one kilometre in diameter likely to cross the Earth's orbit within 10 years.

Congress has requested such studies before. Most recently, the *Spaceguard Survey*, published in 1992, called for a global network of six 2.5-metre telescopes to search the sky for 20 years, at a cost of \$250 million. But the recommendation was ignored, partly because of its price tag.

Two developments have changed the climate since then. One was last month's impact of Comet Shoemaker-Levy with Jupiter, which brought the reality of cosmic collisions home to the general public. The second has been the willingness of the US Defense Department to share observational facilities that could assist in such a search.

The military already uses a network of upward-looking telescopes called GEODSS (Ground-based Electro-Optical Deep Space Surveillance) to track satellites and debris in Earth orbit. Downward-looking spy satellites have for years been recording meteor flashes in the atmosphere, yielding information on objects that collide with Earth.

Gregory Canavan of the Los Alamos National Laboratory also points out that the military's experience in using instruments equipped with charge-coupled devices for long-term, systematic sky searches could be transferred to an automated asteroid watch.

Congress has directed the Department of Defense to work with NASA in developing the search plan, and Pentagon involvement in the search could help ease the funding burden on the space agency. International participation, which Congress has also called for, could bring NASA's costs down even further.

Some have criticized the idea of using nuclear weapons to shoot down or deflect celestial hazards (see *Nature* 370, 65; 1994). But few argue against a programme designed to search for them. Last week, NASA named an eight-member committee to come up with a workable plan. It is headed by Eugene Shoemaker of the Lowell Observatory — co-discoverer of the comet that hit Jupiter last month.

Tony Reichardt

The far Barcoo where they eat nardoo

To carry me westward ho, my boys,
that's where the cattle stray,
On the far Barcoo where they eat nardoo,
a thousand mile away.¹

SIR — Earl and McCleary² provide a perceptive explanation for the illness which, together with starvation, caused the deaths of Burke and Wills. It is a convincing proposal that eating nardoo, prepared without leaching, resulted in excessive thiaminase ingestion and the development of peripheral neuropathy in the explorers. King, the survivor of the expedition, may well have had additional symptoms of cerebral beriberi at the time of his rescue.

The authors are wrong in equating 'Barcoo rot' with scurvy. Barcoo rot is an infective dermatosis, caused by diphtheroid organisms, similar to Veld sore or desert sore³. Susceptibility to the infection may be increased by vitamin deficiency, but the deficiency is predominantly niacin deficiency, resulting in pellagra, rather than lack of ascorbic acid. Albert Gaston⁴ described the disease in miners from Coolgardie (Western Australia): "By continually living on preserved food with a lack of vegetables, the miners' health began to suffer, and a great number of them suffered from Barcoo rot. The skin of the hands and arms became brittle and tender and the slightest scratch soon became a dry hard sore which refused to heal." A more detailed clinical description is given in ref. 3.

Another, entirely different disease of the same period is referred to as Barcoo fever, the symptoms of which were severe nausea, vomiting, often associated with constipation, fever, myalgia, lethargy and possible death. It is probable that this disease was due to ingestion of water containing toxins of cyanobacteria⁵, which were present and produced toxicity in some water sources well before the development of European settlement⁶.

The members of the rear party, who perished on their delayed and futile attempt to move up to the base camp on Cooper Creek, were not suffering from vitamin deficiency and had not eaten nardoo. They died of inanition camped beside the few remaining sources of water between the Darling River and Cooper Creek. Their symptoms were those of hepatotoxicity, and the probable cause was cyanobacterial toxins from contaminated water⁷. They died of 'Barcoo fever', although the disease was probably not known and certainly not named at the time.

The Barcoo River rises in central Queensland and flows south and west for 500 km to join the Thomson River to form the waters of Coopers Creek, beside which Burke and Wills died. They and the

creek will remain in Australian history; the Barcoo and nardoo are part of that account.

John Hayman

*Department of Pathology,
Box Hill Hospital,
Box Hill, Melbourne,
Victoria 3128, Australia*

1. Anon. *The Penguin Book of Australian Ballads* (ed. Ward, R.) 130–131 (Penguin Books, Melbourne, 1964).
2. Earl, J. W. & McCleary, B. V. *Nature* **368**, 683–684 (1994).
3. Ciento, R. *Tropical Diseases in Australia* (2nd edn), 344–347 (W. R. Smith & Paterson, Brisbane, 1942).
4. Gaston, A. *Coolgardie Gold*, 143 (A. H. Stockwell, London, c.1937).
5. Hayman, J. *Med. J. Aust.* **157**, 794–796 (1992).
6. Francis G. *Nature* **18**, 11–12 (1878).
7. Falconer, I. R., Beresford A. M., & Runnegar M. T. C. *Med. J. Aust.* **1**, 511–514 (1983).

No treason

SIR — "Insubstantial charges of treason" (*Nature* **368**, 779; 1994) brings needed scepticism to charges that Enrico Fermi, Robert Oppenheimer, Leo Szilard and others passed nuclear weapons secrets to the Soviet Union. At the same time, however, it perpetuates a common misconception by repeated use of the word "treason" to describe the alleged spying. The US Constitution specifically defines treason as giving aid and comfort to enemies of the United States. The Soviet Union at the time of the Manhattan Project was an ally of the United States and Great Britain in their war against Nazi Germany. The alleged actions of Fermi *et al.* are serious enough to be called by their proper name — espionage — without using an inflammatory term that distorts the historical record.

Curt Covey

*Lawrence Livermore National Laboratory,
Global Climate Research Division,
Mail Code L-264,
PO Box 808,
Livermore, California 94550, USA*

Return of the native

SIR — As a Mexican scientist I should like to comment on the Mondragón syndrome of bad timing (*Nature* **368**, 793; 1994).

Like Alfonso Mondragón, I had the opportunity to pursue postgraduate studies as well as postdoctoral positions in Europe and the United States with possibilities of staying in the United States. But I was convinced that I had to return; I learned from some teachers how science could be done with low resources and the importance of building a scientific community in this developing country. In 1980 I acquired a research associate position at the Instituto de Investigaciones Biomédicas in UNAM. The decade of the 1980s

was very hard indeed, especially for young scientists starting their careers; one needed to ask what experiments could be done with existing resources and not the other way around, and equipment and reagents had to be shared. Much time was (and still is) devoted to teaching and bureaucracy. Not only did people not want to return to Mexico, but many already here left.

We helped to create the Center of Genetic Engineering (now the Institute of Biotechnology) and, looking back, although it has been hard and demanding, we feel satisfied. The institute is productive, we have nurtured several young scientists and although our scientific papers might not be as many as if we had not returned they are equally relevant.

What would have happened if all young scientists in the 1980s had left? In the words of a former US president they asked "not what can the country do for me but what can I do for my country?"

Yes, we do need more scientists but only those willing to work hard to overcome existing problems and not those who feel they deserve special privileges just because they 'sacrifice' themselves by returning.

Patricia Joseph-Brayo

*Instituto de Biotecnología,
Universidad Nacional
Autónoma de México,
Apdo 510–3, Cuernavaca,
Morelos 62271, Mexico*

Animal research

SIR — William Dantzler *et al.* and Susan Paris (*Nature* **369**, 9–10; 1994) rule out any discussions about the use of animals in research with those who pursue abolition, preferring instead to discuss its continuation only through espousing a policy of humane care.

While the ultimate 'goal' of Advocates for Animals remains the total abolition of animals used in research, the society does at the same time seek humane care for animals used in laboratories today. In Britain such a stance has not prevented Advocates for Animals from taking part in meaningful and 'rational' discussions nor hindered it from finding 'common ground' with the scientific community. Only through dialogue, debate and respect for each other's views will this whole controversial issue of the use of animals in research move forward.

Les Ward

*Advocates for Animals,
10 Queensferry Street,
Edinburgh EH2 4PG, UK*

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

Preparing for emerging infections

George A. Gellert

New infectious diseases continue to emerge, yet there is no clear strategy for managing them. A model response should be devised in the light of past events such as the recent US outbreak of a previously unknown hantavirus.

THE era of communicable diseases is hardly over. As well as AIDS, Legionnaire's disease and Lyme disease, other recent emerging infections include *Escherichia coli* O157:H7, cryptosporidiosis, multiple drug-resistant pneumococcus, *Helicobacter pylori* and vancomycin-resistant enterococcus. As human contact with vectors of potentially lethal pathogens increases¹, a US Institute of Medicine report² has emphasized the need for improved surveillance, basic research and laboratory diagnosis for such diseases.

In the spring of 1993, several cases of a severe respiratory disease were reported to the health departments of the Four Corner US states (Arizona, Colorado, New Mexico and Utah). Laboratory evidence implicated a previously unknown rodent-borne hantavirus, and the disease was termed hantavirus pulmonary syndrome (HPS)^{3,4}. By the end of the year, 53 cases of HPS had been reported in 14 states; 32 patients have died. The disease has affected mainly young healthy adults (around 31 years old), nearly half of them American Indians⁵. Transmission seems to be associated with exposure to rodent excreta in aerosol form, particularly that of deer mice. Person-to-person transmission has not been documented.

A team of epidemiologists, clinicians, pathologists and laboratory scientists was established to investigate the outbreak. As the state epidemiologist for Arizona during the outbreak, I led the investigation in the second most heavily affected area. Here I focus on the first four weeks of the investigation, and outline the main operational problems encountered. On the basis of this (and other) experience, I suggest that leading public-health agencies should now devise a model response for dealing with emerging infectious diseases in the future. Although the hantavirus outbreak was restricted to the United States, the recommendations below are applicable internationally.

The initial outbreak involved four state health departments, the Navajo Nation Division of Health, the federal Centers for Disease Control and Prevention (CDC) and the Indian Health Service. Although these were the key decision-making institutions, other organizations and people participated in the investigation, including university medical centres, county health departments, coroners, hospitals and community physicians. Much time and energy however were devoted to defin-

ing and implementing often rudimentary management procedures. Because strategies were developed as needed and in isolation, the response as a whole was inefficient and could easily have become a medical and public policy debacle.

The state and local health departments in the United States do not have a standard protocol to follow when faced with an outbreak of a 'new' infectious disease;

hospitals and local health departments, and mobilize staff and resources.

Although the Institute of Medicine's recommendations² to strengthen state and federal surveillance and to develop national and international databases for infectious diseases are commendable, structures to ensure adequate local communication of surveillance data among overlapping jurisdictions, where detec-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Navajo Indians were affected disproportionately by the hantavirus.

rather, they adopt a narrow technocratic view that holds that if the science is good, then so too will be the response. Without a cure or vaccine, however, rapid and effective organizational responses are vital for disease control.

There are more than 3,000 local and state health jurisdictions in the United States that conduct surveillance for communicable diseases. But new infectious agents are unlikely to appear within the confines of a single jurisdiction. It is therefore important to increase local awareness of possible outbreaks and to ensure early communication among jurisdictions. This did not occur during the HPS outbreak. Southwestern states are sparsely populated, and medical referral patterns routinely cross state boundaries. For example, people in northeastern Arizona usually refer to centres in New Mexico. Poor communication between the states meant that initial Arizonan cases were identified by health authorities in New Mexico at least two weeks before Arizonan authorities knew about them, thus delaying efforts in Arizona to heighten surveillance, communicate with

tion often first occurs, are needed now.

Lacking joint experience in crisis coordination as well as the frameworks for collaboration, state health bureaucracies differed in their priorities, styles, objectives and commitments during the HPS outbreak. There was a need for strong leadership to assess, arbitrate and resolve these differences. Federal leadership is in fact needed. In the United States, the CDC is scientifically, organizationally and politically the institution best suited to assume this role. The CDC can transcend the parochial perspectives of local jurisdictions and promote standards based on state-of-the-art science. But operational leadership has proven difficult for the CDC to assume: the CDC becomes involved in state jurisdictions at the invitation of the state epidemiologist, and there is no legislative mandate that authorizes the CDC to assume such leadership. Legislation should allow for an expansion of the CDC's range of political authority.

A further impediment during the HPS outbreak was caused by concerns over data ownership, intellectual property rights and publication credit. For effective

interagency collaboration and cooperation, therefore, there should be assurances that these issues will be openly and fairly discussed at a later date. Local and federal scientists often exist in an uneasy relationship of mutual need. Although the CDC is the leading national centre for epidemiology and laboratory methods in public health, local institutions are at the front line of the actual outbreaks. Federal teams have been known to take over a local investigation, draw heavily on local resources and depart with data that they later publish as their own. Local practitioners, on the other hand, may not appreciate the constraints on time and resources under which their federal counterparts operate; it is costly to send a federal team to investigate a local outbreak and certain expectations have to be met. And nearly all the teams want to be viewed by the public and their peers as the lead investigative group. Many of these differences could be resolved if there were a systematic way to share credit.

Communication among investigation teams needs to be improved. During the HPS outbreak, daily teleconferences were organized by the CDC with participants from across the United States. The teleconferences were successful in disseminating information about laboratory findings and biosafety recommendations for example, but were ineffective for policy formulation. Participation was patchy, perhaps because investigators were too busy with actual operations, and some organizations were not represented in key discussions. A system for internal communication needs to be established that involves all participants and provides a mechanism for establishing an advance agenda and for giving priority to planning and policy discussions.

Communication with the media and the public is equally important. A new infectious agent is bound to attract a great deal of media attention, and fears and anxieties may result in overly sensational and inaccurate reporting. During the HPS outbreak, the style and extent of media coverage varied among states, partly reflecting the number of deaths in different jurisdictions. But there was no proper coordination of the content and timing of press releases and conferences. Similarly, although recommendations on clinical management for physicians and emergency-room workers were developed, as were occupational safety precautions for health professionals and coroners, these guidelines were not uniformly disseminated across all jurisdictions. Greater federal political authority in the management of outbreaks, and the advent of new technologies such as electronic networks, may offer opportunities for improved communication.

Relative risks need to be conveyed to the public more effectively. In Arizona,

more people die every week from alcohol-related motor-vehicle injuries than those that died from HPS in the first six months of the outbreak. If the gap between public complacency and hysteria is bridged, then people will know how best to alter their behaviour, lifestyle or environment to prevent infection.

Effective media management can also reduce public discrimination of people affected by the disease. The HPS outbreak occurred in a rural area populated predominantly by American Indians; outdoor activities may have increased their exposure to rodents carrying the virus⁵. Public health agencies failed to anticipate that these American Indians would be discriminated against; as a result of what was dubbed "Navajo disease", restaurants refused Indians service, Navajo children were turned away from Los Angeles during a school trip and a travel restriction to affected areas was called for by the media. Public health authorities initially may have been unable to dismiss the possibility of person-to-person transmission, but the persistence of discrimination suggests that they could have done more to prepare for and counter unfair publicity. In low-income communities where tourism is an important source of revenue, discrimination can be as great a threat to public health as the pathogen itself. Vulnerable groups should therefore be identified in advance if possible, and social and behavioural scientists could be called upon to help manage public perceptions.

It is usual practice during the early stage of an outbreak to establish a broad definition of the disease, however inaccurate. With a new pathogen, the case definition must be especially wide to ensure that the full spectrum of disease is detected through surveillance; this strategy was followed in the HPS outbreak. But HPS culminates in adult respiratory distress syndrome (ARDS), which has an annual US incidence of 50,000–150,000 cases in which the cause is unknown. As suspect cases of HPS were excluded in ever greater numbers, it became clear that the initial case definition of HPS was too inclusive. Refinement of the early definition based on epidemiological and clinical criteria depended on laboratory diagnosis. With HPS, laboratory diagnosis was available rapidly because CDC virologists had previous experience with hantaviruses (with the next new infectious agent we may not be so lucky). Yet it was difficult to encourage investigators to narrow the case definition; they were more concerned with apparently more pressing operational matters. Continual reassessment of the case definition is essential for assessing the validity of assumptions and decisions, reviewing strategy and informing the public about the changing magnitude of the outbreak.

Until the pathogen is isolated, clinicians

and epidemiologists practice a speculative science. A central laboratory will often be established to lead in the identification, as the CDC did for HPS. Laboratory diagnostic technologies vary considerably in sophistication and availability. Complex technologies usually have to remain in the central laboratory but others, particularly blood tests, can be transferred to state and local jurisdictions. Rather than continually collecting and transporting specimens across great distances, however, local laboratory capacities and resources could be increased. The feasibility of rapidly equipping local laboratories, and the problems of quality control and variation among institutions, are important issues to balance against possible transfer.

Finally, in a public health crisis, investigators tend to focus exclusively on the short-term issues that immediately confront them. After a while, they become accustomed to planning around narrowly focused issues. Long-term planning for managing the outbreak may thus be excessively delayed. There needs to be a way of assessing when to shift from a crisis mode to long-term planning regarding resource use, epidemiological monitoring and intervention, and internal and public communication. Health officials should consider when epidemiological investigation, experimental treatment protocols and public health education can become less reactive and more routine.

Faced with an outbreak of a new infectious disease, investigators would benefit from the collective wisdom of individuals and agencies who have journeyed into similar lands before. In the United States, the CDC, and, internationally, the World Health Organization should review outbreaks of emerging infections and develop a model response for health agencies involved in investigation and control. Such a process should allow for the fact that emerging infections are not only scientific and medical dilemmas, but also public policy and management problems. When our organizational and management skills and capabilities are as sophisticated, rigorous and well-defined as our methods of scientific investigation, we will have moved far towards meeting the challenge posed by the emergence of new agents of communicable disease. □

George A. Gellert is at the Project HOPE Health Sciences Education Center, Millwood, Virginia 22646, USA.

1. Kilbourne, E. D. *J. Am. med. Ass.* **264**, 68–70 (1990).
2. Lederberg, J., Shope, R. E. & Oaks, S. C. (eds) *Emerging Infections: Microbial Threats to Health in the United States* (Institute of Medicine/National Academy Press, Washington, DC, 1992).
3. Hughes, J. M., Peters, C. J., Cohen, M. L. & Mahy, B. W. J. *Science* **262**, 850–851 (1993).
4. Duchin, J. S., Koster, F. T., Peters, C. J. and the Hantavirus Study Group *New Engl. J. Med.* **330**, 949–955 (1994).
5. Centers for Disease Control and Prevention *Morb. Mortal. wklly Rep.* **4**, 441–443, 477–479, 816–820 (1993); **43**, 45–48 (1994).

Plasma dust as model crystals

Dust particles embedded in ionized plasma can be made to assemble into crystals of a kind, offering new routes with possible astrophysical implications to the study of the crystalline state.

THAT physicists are the masters of artifice in the experiments they design is widely recognized, but sometimes their tricks quite take the breath away. Most of what follows is about an experiment of outstanding ingenuity which, despite the electronic equipment no doubt accompanying it, is also simple. The word 'elegant' is overworked, but in this case it applies.

First, a few words to explain why those concerned have bothered. The crystalline state has been amply studied in the past century or so, leading (among other things) to the modern electronics industry. But nature has been niggardly with the exemplars of the crystalline state it offers. In most crystals, atoms are separated by no more than a fraction of a nanometre. In ionic crystals such as sodium chloride, interatomic distances are no different from those in molecules. In metals, they are even smaller.

To the extent that the materials found useful in the real world are similar, it may be held that it does not matter very much that most studies of the crystalline state have been carried out with crystals in which the atoms are tightly packed. But that overlooks the need to test certain general properties of the crystalline state over a much wider range of parameters than in natural or even synthetic crystals.

How generally valid, for example, is the rule that the crystalline state becomes a liquid in a phase transition marked by the accumulation of disorder in an otherwise regular lattice? Or are the dislocations (which are, in a sense, ordered patterns of disorder) that reduce the strength of real solids below their theoretical strength also found in other kinds of crystals?

Such questions explain the eagerness with which solid-state physicists pounce on every newly discovered system in which there is an ordered array of entities in case it will be a good model of a real crystal. There are now several macroscopic systems that will serve. One is the now-familiar raft of soap bubbles on a liquid surface; that, at least, can simulate dislocations visually. More generally, Langmuir-Blodgett films consisting of unimolecular layers of some material on an immiscible liquid will simulate the melting transition, as will simple molecules adsorbed on a clean metal surface, but these are two-dimensional systems in which the melting transition is not a first-order transition (meaning that there is no latent heat).

The same is true of the 'electron solids' found in thin layers of semiconductor material, where the transition from a solid to a gas is accompanied by a marked change of electrical properties but little else. Other models of the crystalline state include liquid crystals in which weakly interacting molecules are ordered relatively to each other both in space and orientation, and whose utility (in liquid displays) rests on the ease with which the state of order can be changed by a weak electrical field, and colloid solutions, where interparticle forces are usually too weak to make them good models of a crystal.

So why not set out to build another model? That is what a group of six people, five from Germany and one from the University of Iowa in the United States, have now done (Thomas, H. *et al. Phys. Rev. Lett.* **73**, 652; 1994). They call their crystal a 'plasma crystal' formed by the mutual electrostatic repulsion of particles of dust held within an ionized plasma. The authors note the obvious relevance of their study to what may happen in clouds of interstellar dust, which are also 'dusty plasmas', as well as to the practical problems of the contamination of newly formed surfaces by dust in microelectronics fabrication.

How to simulate such a crystal in the laboratory? The plasma is made from argon gas enclosed in a small vacuum chamber and excited by a radio-frequency discharge. The chamber also contains two electrodes, a lower circular electrode 8 cm in diameter and an upper electrode in the form of an annulus — a 10-cm circle in which a 3-cm hole has been cut. Both electrodes are arranged horizontally and are 2 cm apart.

For dust, Thomas *et al.* have used 7 μm plastic spheres, which are dusted into the plasma chamber through the hole in the upper electrode. In the conditions of the experiment, the plastic spheres form a circular cloud some 3 cm in diameter. The positions of the dust particles are made visible by means of a fan-shaped laser beam (made by putting a cylindrical lens in front of a laser), when the scattered light can be seen (even by the naked eye, it appears) through the central hole in the upper electrode and also recorded by means of a charge-coupled device for numerical analysis.

What happens? In one run of the experiment described in detail, the dust grains are arranged in 18 layers all parallel with the horizontal electrodes. The positions of the

particles within each layer are determined simply by shifting the laser beam vertically. Even to the eye, there is a high degree of regularity in the arrangement of the particles in a plane; many points of light form nearly perfect squares, others form trapezia with the same dimensions, but others appear to be missing altogether. The authors estimate that the inter-particle spacing is 250 μm , or 0.25 mm.

The explanation is straightforward. One clue is that in the experiment described, the lower electrode acquires a negative electrical bias of 14 V, which is enough to levitate the dust particles. The authors explain that the electrical charge on a particular dust particle will be determined by the competitive attachment of electrons and ions to its surface, but, in this experiment, the electrons win, the dust particles are negatively charged and, although they are presumably surrounded by partially neutralizing sheaths of argon ions (as in the Debye-Hückel picture of ions in aqueous solution), they repel each other electrostatically.

How good is the model? There are now well-known techniques for telling whether an array of dots in a two-dimensional plane embodies a degree of order: join each pair of neighbour-points with a straight line, bisect that line and draw a line at right-angles through the bisection point until it meets others like itself. The result is that each point is embedded in what is called a Voronoi cell. The most telling proof that Thomas *et al.* have indeed made a 'dusty plasma crystal' is that their distributions of particles in horizontal planes are strikingly more regular than would be obtained with a random distribution of points in a plane.

What happens next? The obvious lack in this first account is that there is no investigation of the three-dimensional properties of the dusty crystal. And a first attempt at observing the melting transitions (by increasing the radio-frequency power) seems to have been swamped by the large motions imparted to all the dust particles, but that test will ultimately hang on a way of feeding energy directly into the vibrational modes of the dusty crystal.

What the authors themselves would like to do is to carry out their experiment in 'microgravity' conditions. On the face of things, it is a more deserving candidate for room in some future Space Shuttle flight than most of the others so far suggested.

John Maddox

Moving forward noisily

S. Leibler

INSPIRED by biological systems or motivated by biological applications, physicists are currently creating and studying new devices acting on submicrometre scales. The latest example is the microscopic structure presented by Rousselet, Salome, Ajdari and Prost on page 446 of this issue¹. It consists of an array of 'sawtooth'-shaped electrodes which can be made to propel small charged colloidal particles by alternately switching on and off the electric potential. This structure can be viewed either as a device inspired by the action of motor enzymes, such as kinesin or myosin, or as a first step towards new separation methods for biological macromolecules, chromosomes or viruses.

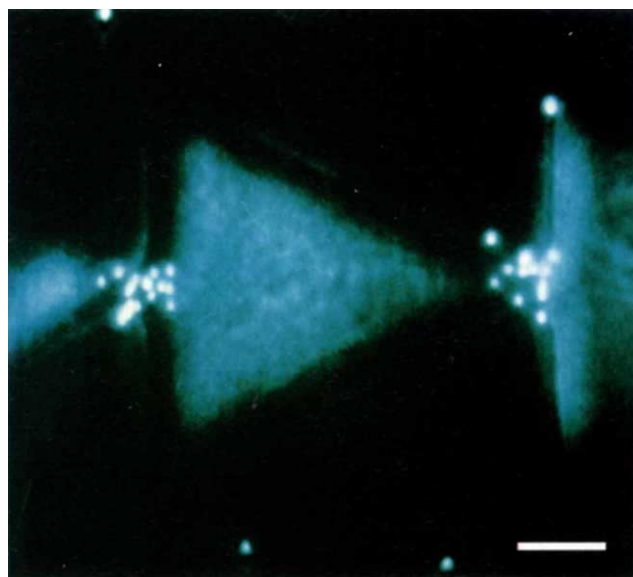
Let us take the possible applications first. This and similar devices are fabricated by the microlithography techniques developed mainly for semiconductor electronics, which allow one to create well defined structures with submicrometre dimensions. Particles fitting in and moving inside structures of such small sizes are inevitably subject to thermal noise: they are 'brownian particles'. To master their behaviour one has to understand the stochastic processes involved in the movement. Rousselet *et al.* show that not only can one successfully deal with the noisy thermal environment on submicrometre scales, but also, by adding to it some extra component, possibly accomplish useful tasks.

Although the techniques allowing the creation of the present device are relatively new, the underlying ideas are old. A hundred years ago Pierre Curie defined² the symmetry principle on which this device is based. It implies that periodic structures with spatial asymmetry may act as 'ratchets' for brownian particles in the presence of dissipation or other source of time-reversal symmetry breaking. By a ratchet we mean here a system which moves the particles with non-zero macroscopic velocity without any macroscopic forces or field gradients. Such a ratchet is thus different from, for instance, a simple electrophoretic machine in which the moving particles are subject to a macroscopic electric force.

Two years ago, Ajdari and Prost³, working in Pierre Curie's old institution in Paris, pushed his ideas further, showing in theory how a simple microstructure could act as a ratchet or a separation device.

Imagine brownian particles moving in a sawtooth potential, $U(x)$, whose 'teeth' are periodic but asymmetric. If the particles are subject only to thermal equilibrium noise, then their movement is not macroscopically biased in one of the directions along the x -axis and the sawtooth potential cannot act as a ratchet for brownian particles⁴.

But if one now constantly switches the potential $U(x)$ on and off, then the particles can move in one direction with



Trickle down — tiny colloidal spheres trapped at the narrow necks between 'Christmas tree' electrodes. Each on-off cycle of electric field produces a net particle motion from left to right. Scale bar, 10 μm . (Photo courtesy of J. Prost.)

non-zero macroscopic velocity. During the period in which the potential is switched off, the particles diffuse freely and their distribution spreads along the x -axis. Although the diffusion is symmetrical, the potential U is not, so when it is switched back on it pulls more particles in one direction than in the other.

The calculation shows that in the simplest case, in which the potential is switched periodically, the macroscopic flux depends exponentially on the diffusion coefficient D of the particle (at least in the case of slow diffusion and thus low D). This is particularly promising from the point of view of using such systems to separate particles of different diffusion constant, for instance because of differences in size.

Ajdari and Prost have now gone beyond this purely theoretical work: together with their colleagues from Bordeaux¹ they have built a microelectrode system based on exactly the same principle of periodically appearing and disappearing sawtooth electrostatic potential, and they have demonstrated that it does indeed

move forward small colloidal particles of sizes varying between 0.1 and 5 μm (see figure). Although inevitable complications of the real system, such as sticking of the particles and non-isotropic diffusion, prevent quantitative comparison with theoretical predictions, the data agree at least semi-quantitatively with the simple theory.

The success of this exemplary interplay between theory and experiment does not yet imply the success of future applications. Although the present work clearly shows that biased movement and separation can be achieved, it remains to be seen whether the method can give better results than more traditional techniques, such as electrophoresis, in which macroscopic gradients or forces are directly applied, especially as a new class of such techniques has been introduced which also takes advantage of the microlithographic technology. For instance, one can perform electrophoresis of DNA molecules in a submicrometre 'maze' structured in a silicon wafer⁵. Further development of electrophoresis in such well controlled geometries might lead to separation of relatively large brownian particles (megabase pieces of DNA or even whole cells) for which the usual gel methods are inefficient.

Periodic switching of the sawtooth potential is not the only way of introducing the time correlations necessary for macroscopic movement — a series of theoretical papers has more recently explored other

possible ratchets. Magnasco⁶ looked at the case in which the particles are subject to a periodically varying macroscopic force, $F(t)$. Strictly speaking this is not a ratchet in the sense defined above, but it is analogous to it, provided that the time average of $F(t)$ is zero. Analysis of this model showed that macroscopic movement takes place for various functional forms of $F(t)$.

Similar results can be obtained if one simply modulates the height of the teeth of the potential. This can be done, as before³, by switching between two different 'sawtooth' potentials⁷, or equivalently, by assuming that a brownian particle can be in two different internal states, each of which feels a different sawtooth potential (ref. 8; and G. Oster, C. S. Peskin and G. B. Ermentrout, personal communication). Instead of introducing the time correlations in the potential $U(x)$ itself, one introduces them in the transitions between two internal states of the particle. Out-of-equilibrium, correlated noise (so-called 'coloured noise') is again necessary: if the transition rates corres-

pond to thermal equilibrium ('white noise') then the particle cannot move on average in one direction⁹. In the presence of coloured noise the average flux depends strongly on the noise amplitude, in some cases showing a maximum for a well defined value⁷.

Earlier this year Doering *et al.*¹⁰ investigated how the average velocity of the particles depends on the characteristics of the out-of-equilibrium noise. Their combination of numerical and analytical results shows that not only the magnitude of the flux but also its sign depends on the nature of the noise. This means that for a given sawtooth potential the particles may be moved in either direction by simply changing the characteristics of time correlations in the nonthermal component of the noise.

Most of this theoretical activity was inspired not by Curie's old results but by developments in the study of motor enzymes¹¹. Molecular motors, such as kinesins or myosins, move on protein fibres many micrometres long, and can transport objects such as vesicles or chromosomes within the cell, participate in cell locomotion or accomplish other types of useful mechanical work. The energetics of the action of molecular motors has long been known¹²: hydrolysis of bound nucleotides such as ATP provides the necessary energy. But the dynamics of this process is less well understood, although the crystal structure of muscle myosin has now been established¹³ and the action of single motors has been observed with impressive nanometre resolution^{14,15}. Not surprisingly, experimental data like these have inspired theorists to new efforts.

Motor enzymes and ratchets are indeed analogous in many respects. Both work in the presence of thermal noise and transport brownian particles; the ATP hydrolysis which drives a molecular motor is local, not involving gradients or forces on micrometre scales; the fibre on which a motor enzyme moves shows periodicity and spatial asymmetry because of the nature of its protein components; the binding sites for a motor on a fibre are analogous to the minima of a sawtooth potential; and the movement of a motor is directional, but different motors may move in different directions on the same fibre.

It is tempting to take the extra step and argue that a ratchet and a motor enzyme work in exactly the same way. Then one can try to compare quantitatively their characteristics⁷. Unfortunately, one cannot then get anything beyond a simple dimensional analysis. The immediate problem is which of the ratchet models is the most appropriate. For instance, can the hydrolysis of ATP be considered as a source of a 'coloured', out-of-equilibrium noise which moves the motor in a fixed fibre potential? Or is it better to say that

the motor enzyme can be found in two different states and ATP hydrolysis introduces nonthermal transitions between these states? One could also imagine a motor locally modifying the fibre on which it moves and thus changing the sawtooth potential, or some still more complicated situation.

The point is that all these questions are of microscopic nature and will be answered only by further experimental studies of the details of molecular motors. It is useful to know that simple physical devices can perform tasks analogous to those of biological systems, but this does not mean that the latter are as simple as the former.

Should we even try quantitatively to model the motor enzymes? Well, the effort might be worthwhile provided that one stays on the phenomenological rather than the microscopic level. For instance, *in vitro* motility assays¹⁶ characterizing the action of molecular motors tell us that the average motor velocity depends on ATP concentration, or on the load acting on the motors. Similarly, more quantitative motility assays might establish the relation between the mean hydrolysis rate and motor velocity, or correlations between the statistics of hydrolysis events and fluctuations in their movement. Phenomenological models explaining all these relations could also give some testable predictions, point to the differences in the action of different motors¹⁷ or even suggest some molecular (genetic) modifications which could alter the action of the motors.

Meanwhile, the work of Rousselet and her colleagues provides us with a beautiful example of biologically inspired physics and, in addition, a device with possible applications in biological research. □

S. Leibler is in the Departments of Physics and of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA.

1. Rousselet, J., Salome, L., Ajdari, A. & Prost, J. *Nature* **370**, 446–448 (1994).
2. Curie, P. J. *Phys. III*, 383 (1894).
3. Ajdari, A. & Prost, J. *C.R. Acad. Sci. Paris* **315**, 1635–1639 (1992).
4. Feynman, R. P., Leighton, R. B. & Sands, M. in *The Feynman Lectures in Physics* (Addison-Wesley, Reading, Massachusetts, 1966).
5. Volkmuth, W. D. & Austin, R. H. *Nature* **358**, 600 (1992).
6. Magnasco, M. *Phys. Rev. Lett.* **71**, 1477–1481 (1993).
7. Astumian, R. D. & Bier, M. *Phys. Rev. Lett.* **72**, 1766–1769 (1994).
8. Prost, J., Chauvin, J.-F., Peliti, L. & Ajdari, A. *Phys. Rev. Lett.* **72**, 2652–2655 (1994).
9. Vale, R. & Oosawa, F. *Adv. Biophys.* **26**, 97–134 (1990).
10. Doering, C. R., Horsthemke, W. & Riordan, J. *Phys. Rev. Lett.* **72**, 2984–2987 (1994).
11. Maddox, J. *Nature* **369**, 181 (1994).
12. Hill, T. L. *Thermodynamics for Chemists and Biologists* (Addison-Wesley, Reading, Massachusetts, 1968).
13. Rayment, I. *et al. Science* **261**, 50–58 (1993).
14. Svoboda, K., Schmidt, C. F., Schnapp, B. J. & Block, S. M. *Nature* **365**, 721–727 (1993).
15. Finer, J. T., Simmons, R. M. & Spudis, J. A. *Nature* **368**, 113 (1994).
16. Scholley, J. M. *Motility Assays for Motor Proteins* (Academic, New York, 1993).
17. Leibler, S. & Huse, D. A. *J. Cell Biol.* **121**, 1357 (1993).

RÉSUMÉ

Radical solution

At last, some cheering news for smokers. Dosing up with vitamin C can protect against one of smoking's nastier side effects — the clumping of white blood cells on vessel walls, which can trigger cardiovascular and pulmonary disease (H. A. Lehr, B. Frei and K. E. Arfors, *Proc. natn. Acad. Sci. U.S.A.* **91**, 7688–7692; 1994). Vitamin C, which is a water-soluble vitamin, acts as an antioxidant, disarming poisonous oxygen radicals from cigarette smoke that cause the white blood cells to become sticky. But the water-insoluble vitamin E, another antioxidant that works by a different mechanism, does not confer the same protection. The secret is probably that the oxygen species responsible for the effect on a smoker's leukocytes are more soluble in blood plasma.

Old crocs

CROCODILES today like their climate balmy — 25 to 35 °C is their preference, and anything above 39 °C or below about 4 °C is inhospitable. So P. J. Markwick has been comparing the distributions of ancient crocodilians with their modern counterpart *Alligator mississippiensis* and drawing inferences about continental temperatures across North America in the Cenozoic era (*Geology* **22**, 613–616; 1994). Fossil crocodilians from the Eocene (38–55 million years ago) are widespread, but by the late Oligocene the populations had apparently shrunk to little more than modern-day Florida. The Miocene again was crocodile-friendly, but the Pleistocene and early Holocene were emphatically not. The crocodilian low-down on climate, part of a larger compilation of vertebrate records, thus backs up floral and modelling evidence for climes milder than today's during both Eocene and Miocene.

High charges

THE simplest of all atomic systems are those that, like the hydrogen atom, have just one electron and so avoid the complexities of multielectron interactions. And the higher the atomic number of an element, the greater the contributions of quantum electrodynamics and relativity to the atomic energy levels. The ideal testbed for theoretical models, then, is a heavy ion with every electron but one stripped away — such as the high-velocity U^{91+} ions produced in a storage ring last year (T. Stohler *et al. Phys. Rev. Lett.* **71**, 2184; 1993). Now R. E. Marrs and co-workers have gone further by trapping single-electron U^{91+} , and even a few naked U^{92+} ions, in a compressed electron beam (*Phys. Rev. Lett.* **72**, 4082–4085; 1994). With the ions more or less stationary, they were able to derive cross-sections for electron impact ionization, and suggest that current theoretical estimates are too low.

Exploring signalling pathways

N. H. Brown and D. A. Hartley

A PLETHORA of new data on the development of the fruitfly *Drosophila* was presented by the participants of the latest biennial gathering of Drosophilists in Crete*. One thread running through the maze of diverse developmental paradigms is the importance of signalling pathways in the specification of cell fates.

Much discussion centred on the pathways mediated by secreted factors such as those encoded by *decapentaplegic* (*dpp*, of the transforming growth factor- β (TGF- β) family) and *wingless* (*Dwnt*), and receptor tyrosine kinases such as *sevenless* and the *Drosophila* EGF receptor. Using ectopic expression in clones, K. Basler (Univ. Zurich) showed that *dpp* can generate pattern duplications in both the anterior and posterior compartments of the wing, suggesting that it normally organizes pattern on both sides of the anterior-posterior boundary (see figure). The expression of *dpp* along the wing anterior-posterior boundary depends on *engrailed* and *hedgehog*, which themselves control *wingless* expression at the embryonic parasegmental boundary (*a* in figure). In the embryo, epidermal expression of *dpp* also signals to the developing mesoderm (M. Frasch, Mount Sinai School of Medicine, New York). Here it is required to maintain the expression of one homeobox gene, *tinman*, which in turn activates another, *bagpipe*, which specifies visceral mesoderm.

The words 'morphogen' and 'organizer' always provoke healthy debate with re-

gards to *wingless* signalling. Evidence for signalling over several cell diameters in the developing wing (J. Axelrod, Harvard Medical School) contrasted with short-range effects observed both on *engrailed* expression in the embryo (J. P. Vincent, MRC Lab. of Molecular Biology, Cambridge) and in the specification of ventral cells in the developing leg (S. Cohen, EMBO Lab., Heidelberg). Although the *wingless* receptor(s) remains elusive, candidate *dpp* receptors have been found — *Saxophone* and *thick vein* encode serine/threonine kinase receptors homologous to type-I TGF- β receptors (W. Gelbart, Harvard Univ.; M. Hoffmann, Univ. of Wisconsin). Other possible components of the pathway include the *Mothers against dpp* (*Mad*) and *Medea* (*Med*) genes (see figure).

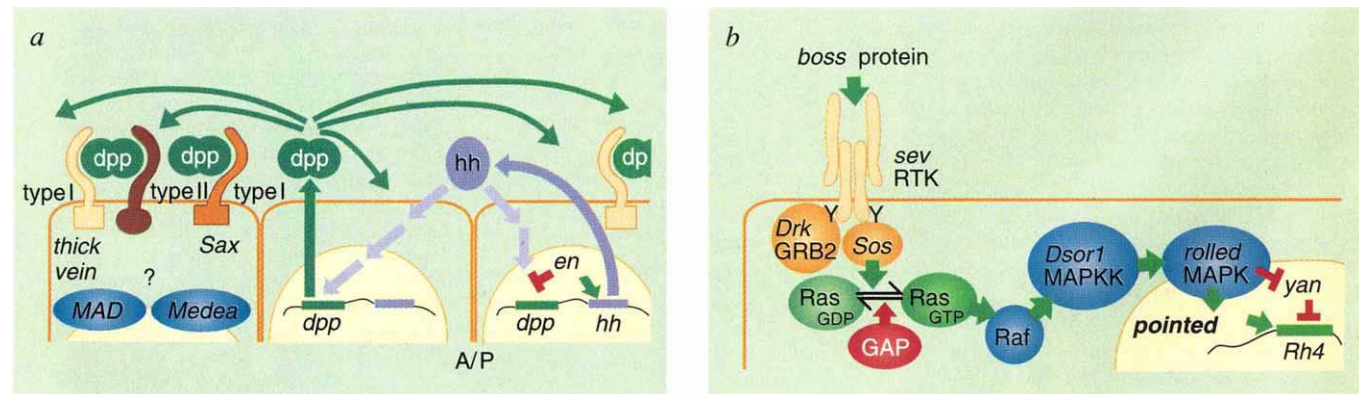
Nuclear targets of receptor tyrosine kinase (RTK) signalling were described. Hitherto, genetic screens have uncovered members of the RTK/*ras* pathway leading to the nuclear MAP kinase (*b* in figure). Potential transcription factor targets of MAP kinase were described that are encoded by the *yan* and *pointed* genes and are related to the ETS-domain proteins (G. Rubin, Univ. California, Berkeley; E. Hafen, Univ. Zurich). Both can be phosphorylated *in vitro* by activated MAP kinase and both require MAP kinase phosphorylation sites for activity *in vivo*. Whereas *yan* acts as a repressor, *pointed* acts as a positive regulator, suggesting that the RTK/*ras* pathway results in the simultaneous activation of an activator and inactivation of a repressor, a situation that would result in tight regulation of

target genes. Consistent with the general requirement of the *ras* pathway in RTK signalling, *pointed* is also a target for EGF receptor signalling during mid-line glia development (C. Klämbt, Univ. Cologne).

The link between membrane receptors and the nucleus was also described for the *Notch* signalling pathway. In cell culture, the *Suppressor of Hairless* gene product appears to move from the cytoplasm to the nucleus in response to the binding of *Notch* to its ligand *Delta* (S. Artavanis-Tsakonas, Yale Univ.). In the eye, the ability of activated *Notch* to block the RTK/*ras* signalling pathway supports the idea that overall function of *Notch* signalling is to block signals that lead to the differentiation of uncommitted cells.

The function and regulation of homeobox genes continues to yield surprises. Generally, it is assumed that the specificity of transcription factors resides in their ability to bind specific DNA sequences. So, for example, deletion of the homeodomain of *fushi terazu* (*ftz*) blocks its ability to bind its target site (H. Krause, Univ. Toronto); remarkably, in the embryo the same deleted form behaves identically to the wild-type protein in several assays. This suggests that for *ftz*, protein-protein interactions are at least as important as DNA-protein interactions in regulating transcription. Specific patterns of *Ultrabithorax* (*Ubx*) expression within the imaginal discs are dependent on a two-step process in the embryo (V. Pirota, Univ. Geneva). In the first step, *Ubx* is activated and repressed primarily by gap genes such as *hunchback*, which demarcates the anterior boundary of *Ubx* in a manner suggesting direct occupancy of promoter sites. This may provide a template for the second step, the stable repression of en-

*EMBO workshop on the Molecular and Developmental Biology of *Drosophila*, Kolymbari, Crete, 19–25 June 1994.



Key signalling pathways in the development of adult *Drosophila* body parts. *a*, Induction of tissue polarity at the anterior-posterior (A/P) compartment boundary in the wing by the TGF- β superfamily member, *decapentaplegic* (*dpp*). Serine/threonine kinase receptors have been found of the type-I form (*thick vein* and *Saxophone*). The exact composition and mechanism of *dpp* receptor in this context is not fully understood. Potential downstream components of the receptor are the *Medea* and *Mothers against dpp* (*Mad*) gene products. Tight regulation of *dpp* expression at the boundary is achieved through *engrailed* (*en*) and *hedgehog* (*hh*). *b*, Induction of the R7 photoreceptor cell through

the *sevenless* receptor tyrosine kinase by its ligand *boss*. Activated *Sevenless* recruits the SH2-SH3 adaptor, *Drk*, and the nucleotide-exchange factor, *Sos*, to the membrane. This leads to activation of *Drosophila* *Ras* and in turn *Drosophila* *Raf*. The cascade continues through the MAP-kinase kinase (MAPKK) encoded by *Dsr1*, and the MAP kinase encoded by *rolled*, which moves into the nucleus to phosphorylate two ETS-like transcriptional regulators directly, activating a positive one (*pointed*) and repressing a negative one (*yan*). This leads to simultaneous activation and derepression of target genes like the rhodopsin *Rh4*.

hancers into 'packaged' chromatin through the action of the *Polycomb* group of genes. This strategy restricts *Ubx* expression in the imaginal discs to the appropriate domain. Indeed, *cis*-acting *Polycomb* response elements in the upstream region of *Ubx* were described which repress the activity of adjacent genes over some distance.

The value of *Drosophila* as a general model for animal development was emphasized by some new findings. The orthologue of the PAX-6 (*smalleye*) gene of mouse is encoded by the fly gene *eyeless*, implying an unforeseen evolutionary relationship between the compound eye and the vertebrate eye (W. Gehring, Univ. of Basel). Screens for tumour-suppressor genes have identified interesting phenotypes (T. Xu, Yale Univ.). Also, the *gypsy* retrotransposon may function as a *bona fide* retrovirus, packaged into infectious virus-like particles (A. Bucheton, CNRS, Gif-sur-Yvette; V. Corces,

Johns Hopkins Univ.). Perhaps humans, with their conspicuous lack of wings, should be considered as underdeveloped flies; however, that icon of homeotic transformation, the four-winged fly, is also incapable of flight. Attempts to correct this deficiency have been undertaken using more mutations in one locus than most geneticists use in a genome (*abx bx³ Ubx^{61d} pbx iab-2*; E. Lewis, Caltech). Alas, this tribute to the hundredth year of homoeosis after Bateson was unsuccessful, but we look forward to the 1996 Crete meeting and a flying four-winged *Drosophila*. □

N. H. Brown is at the Wellcome/CRC Institute, Tennis Court Road, Cambridge CB2 1QR and the Department of Biochemistry, University of Cambridge, Cambridge CB2 1QW, UK. D. A. Hartley is in the Department of Biochemistry, Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK.

PALAEOCLIMATOLOGY

Chill over the Cretaceous

Eric J. Barron

THE mid-Cretaceous period, about 100 million years ago, was far warmer than today and thus should be a good test of how well models of future climate can simulate conditions very different from the present day. Or so we thought. On page 453 of this issue, Sellwood *et al.*¹ challenge the evidence for extreme Cretaceous warmth, suggesting the need to re-evaluate past modelling efforts and to question whether the Cretaceous is a reasonable analogue for global change studies.

Much of the effort so far has been based on an early synthesis of the Cretaceous climate observations² indicating substantially warmer poles (mean annual temperatures above freezing) and small but significant increases in tropical sea surface temperatures. Like these earlier studies, Sellwood and colleagues' conclusions are based on oxygen isotope analyses of planktonic foraminifera, but they examine more locations, including more sites at lower latitudes, for a more specific time interval.

The results suggest that mean annual polar temperatures were near 0 °C, so some glacial ice could have formed, and that tropical sea surface temperatures were similar to modern values. This interpretation is very similar to the 'minimum' plausible warmth scenario of earlier Cretaceous climate reconstructions². If the new results are correct, their reconstruction of Cretaceous temperatures has considerable significance.

First, the nature of tropical climate

change is critical given the importance of the tropics in the global heat engine which drives the atmospheric circulation and the likely sensitivity of tropical life to a temperature change of even a few degrees. Most climate models predict that future tropical surface ocean temperatures will rise by about 2 °C for a doubling of carbon dioxide in the atmosphere. Similar model predictions result for the Cretaceous with higher CO₂ (see ref. 3). If Sellwood and co-workers are right, then either other (possibly unknown) climate forcing factors must have counteracted tropical warming, or our understanding of the climate system, as reflected in our ability to construct numerical climate models, is inadequate.

One well-known limitation of current climate models is inadequate treatment of the heat transport of the oceans. Most simulations of future and past climates are based on models which only consider the energy balance in the upper mixed layer of the ocean, lacking ocean dynamics or ocean heat transport. If the oceans transport more heat polewards, the result could be cooler tropical temperatures³⁻⁵. But we do not know that the oceans can play a different role in poleward heat transport, and other mechanisms may be required.

Second, we should look again at efforts to assess the sensitivity of the climate system to forcing factors such as carbon dioxide. For example, geochemical observations and models estimate that Cretaceous CO₂ concentrations were 4 to 8 times present-day values⁶ — this, in-

deed, is a reason why the Cretaceous is so attractive to climate modellers. Using a 'warmer' estimate of the Cretaceous with warmer tropics and polar temperatures near 10 °C (global average temperature increase of 12 °C relative to the present day) would yield an average climate sensitivity to a carbon dioxide doubling of 3–6 °C. These values are on the upper end of the spectrum of climate sensitivity used in current assessments of the potential impact of human-induced global warming⁷. Of course, this is a crude analysis given that the sensitivity is not linear with respect to carbon dioxide concentration in the atmosphere. In contrast, the 'minimum' estimates of Cretaceous warmth are less than 6 °C, yielding an average climate sensitivity which must be closer to 1.5–3.0 °C for a doubling of carbon dioxide, on the lower end of CO₂ sensitivity estimates. This, of course, depends on the importance of other forcing factors and any uncertainty in the estimates of the level of carbon dioxide in the Cretaceous atmosphere.

The large number of differences between the past and the present day make it unlikely that any past climate can be a true analogue of future climates. This does not diminish the importance of the geological record in providing 'case studies' of global change, yielding valuable insights into climate change and climate sensitivity; but it shows that our ability to assess these critical aspects of the Earth system depend on our ability to reconstruct key past observations.

Sellwood and colleagues' interpretations are likely to be challenged. Do other time slices exhibit similar tropical temperatures or are tropical temperatures capable of change on timescales of millions of years through changes in ocean circulation? Are the limited data available dependent on local environmental conditions such as upwelling and is there substantial longitudinal variability? What is the true difference between surface temperature and the temperature at the depths in which the foraminifera lived? Still, they have undoubtedly given fresh life to the debate on climate and climate change. □

Eric J. Barron is at the Earth System Science Center, The Pennsylvania State University, 248 Deike Building, University Park, Pennsylvania 16802, USA.

1. Sellwood, B. W., Price, G. D. & Valdes, P. J. *Nature* **370**, 453–455 (1994).
2. Barron, E. J. *Earth Sci. Rev.* **18**, 305–338 (1983).
3. Barron, E. J., Peterson, W. H., Pollard, D. & Thompson, S. *Paleoceanography* **8**, 785–798 (1993).
4. Rind, D. & Chandler, M. J. *geophys. Res.* **96**, 7434–7461 (1991).
5. Covey, C. & Thompson, S. L. *Global planet. Change* **1**, 331–341 (1989).
6. Berner, R. A. *Nature* **358**, 114 (1992).
7. Houghton, J. T., Callander, B. A. & Varney, S. K. (eds) *Climate Change 1992: The Supplementary Report to the IPCC Scientific Assessment* (Cambridge Univ. Press, 1992).

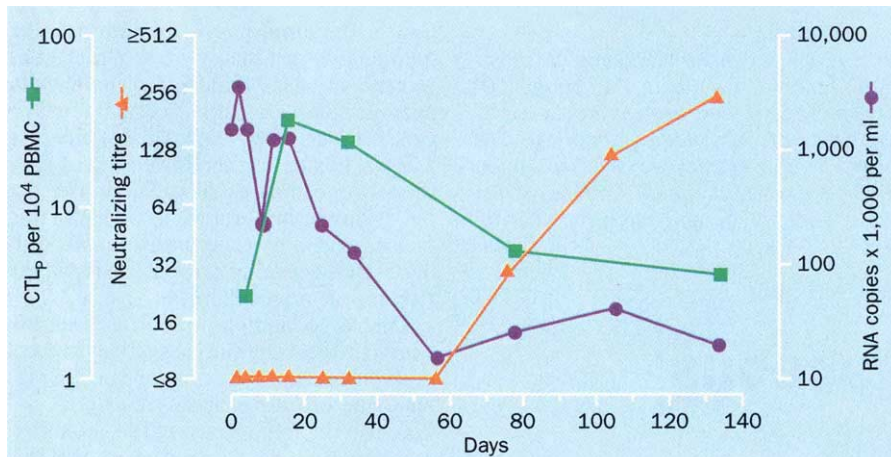
Shutting down HIV

Richard A. Koup and David D. Ho

MANY patients newly infected with the human immunodeficiency virus type 1 (HIV-1) experience an acute, self-limited illness that is characterized by a burst of viraemia followed by rapid, though often incomplete, control of the virus¹. These patients offer a unique opportunity to define the immune responses directed against HIV-1. On page 463 of this issue²,

virus from his plasma had a delayed CTL response⁷. In rhesus monkeys acutely infected with the simian immunodeficiency virus, the CTL response also develops early and is important in controlling the initial viral replication^{8,9}.

In contrast to the CTL response, a detectable neutralizing antibody response to acute HIV-1 infection is substantially



Temporal changes in viraemia, frequency of HIV-1-specific CTL precursors (CTL_p) (per 10⁴ peripheral blood mononuclear cells (PBMC)) and neutralizing antibody response in a patient with acute HIV-1 infection. Plasma viraemia was measured by a commercial branched DNA assay (Chiron); sequential CTL_p frequencies against HIV-1 Pol-expressing cells and neutralizing antibody titres against the autologous virus were determined as previously described.

Pantaleo *et al.* describe the oligoclonal expansions of CD8⁺ T lymphocytes in six patients with primary HIV-1 infection, thereby providing indirect evidence for the involvement of cytotoxic T lymphocytes (CTLs) in the initial immune response against the virus.

Major histocompatibility complex (MHC)-restricted, virus-specific CTLs are important in protective immunity against viral infections. The potent antiviral activity of these cells has been demonstrated both *in vitro*³ and *in vivo*⁴. Now we have evidence pointing to the crucial role of CTLs in the control of acute HIV-1 infection. CTL clones specific for HIV-1 Env protein have been expanded from the blood of patients as early as two days after the onset of acute symptoms^{5,6}. Importantly, these clones specifically recognize sequences expressed by the transmitted virus.

During acute HIV-1 infection, the number of CTL precursors directed against the viral Gag, Pol and Env proteins peaks as viraemia declines rapidly (see figure)⁷. This temporal profile of the CTL response suggests that it is involved in shutting down initial HIV-1 replication, a conclusion that finds support in the anecdotal observation that one acute seroconverter who was unable to eliminate infectious

delayed (see figure)⁷. Antibodies capable of binding a number of epitopes on the viral glycoprotein gp120, however, are detectable almost as early as the CTL response¹⁰. It is formally possible that such antibodies contribute to viral clearance by mechanisms other than neutralization, for example by opsonization and antibody-dependent cellular cytotoxicity.

Now Pantaleo *et al.* observe that the CD8⁺ T-cell expansion during acute HIV-1 infection is oligoclonal². In one subject examined, the expanded T cells appear to represent antigen-specific CTLs, although the level of HIV-1-specific lytic activity exhibited by the expanded cells is far below that expected from a clonal response. The authors suggest, however, that the low level of lysis reflects sequence differences between the strain of HIV-1 that infected the patient and the strain used to test for lytic activity.

The oligoclonal expansion of virus-specific CTLs during acute infection would not be entirely unexpected. Although few studies of V β repertoires during acute viral infections have been reported, perturbations in the V β profiles and CDR3 lengths of the T-cell receptors on CD8⁺ cells are common^{11,12}, and may represent the presence of acute viral infec-

tions in the populations under study. It is also known that specific antigens can stimulate dramatic expansion of CD8⁺ cells expressing certain V β families¹³. In addition, some class-II-restricted 'promiscuous peptides' have been found to stimulate expansion of CD4⁺ cells expressing specific V β families¹⁴. It is therefore possible that an acute viral infection results in oligoclonal CD8⁺ T-cell expansions because immunodominant epitopes stimulate specific V β -expressing CD8⁺ cells. Work in press shows that infection of outbred pigtail macaques with the PBj14 strain of simian immunodeficiency virus results in the acute expansion of CD4⁺ and CD8⁺ T cells expressing V β 7 and V β 14 gene families¹⁵. These expansions may represent a superantigenic property of PBj14, perhaps unrelated to the phenomenon reported by Pantaleo *et al.*

A limited oligoclonal CTL response to acute HIV-1 infection may permit the development of viral variants that escape the initial response, but data are few. In four acute HIV-1 seroconvertors from whom Env-specific CTL clones were isolated, no escape viruses were detected within the first four months after infection^{5,6}. Does the loss of the oligoclonal T-cell expansions after the acute infection syndrome represent clonal exhaustion, clonal shrinkage secondary to antigen clearance, or diversification of the HIV-1-specific response resulting in the dilution of individual clonal responses? Because the CTL response generally remains vigorous and broadly reactive several years into HIV-1 infection, it is unlikely that clonal exhaustion is operative during the acute phase. Further studies with CTL clones from seroconvertors will be needed to address these and other unresolved matters. As for now, the available evidence suggests that the CTL response provides protective immunity in the acute phase of HIV-1 infection. □

Richard A. Koup and David D. Ho are in the Aaron Diamond AIDS Research Center, New York University School of Medicine, 455 First Avenue, 7th Floor, New York, New York 10016, USA.

- Daar, E. S., Moudgil, T., Meyer, R. D. & Ho, D. D. *New Engl. J. Med.* **324**, 961–964 (1991).
- Pantaleo, G. *et al.* *Nature* **370**, 463–467 (1994).
- Tsubota, H. *et al.* *J. exp. Med.* **169**, 1421–1434 (1989).
- Byrne, J. A. & Oldstone, M. B. A. *J. Virol.* **51**, 682 (1984).
- Safrit, J. T., Andrews, C. A., Zhu, T., Ho, D. D. & Koup, R. A. *J. exp. Med.* **179**, 463–472 (1994).
- Safrit, J. T., Lee, A. Y., Andrews, C. A. & Koup, R. A. *J. Immun.* (in the press).
- Koup, R. A. *et al.* *J. Virol.* **68**, 4650–4655 (1994).
- Yasutomi, Y. *et al.* *J. Virol.* **67**, 1707–1711 (1993).
- Reimann, K. A. *et al.* *J. Virol.* **68**, 2362–2370 (1994).
- Moore, J. P., Cao, Y., Ho, D. D. & Koup, R. A. *J. Virol.* **68**, 5142–5155 (1994).
- Posnett, D. N., Sinha, R., Kabak, S. & Russo, C. J. *exp. Med.* **179**, 609–618 (1994).
- Hingorani, R. *et al.* *J. Immun.* **151**, 5762–5769 (1993).
- MacDonald, H. R., Casanova, J.-L., Maryanski, J. L. & Cerottini, J.-C. *J. exp. Med.* **177**, 1487–1492 (1993).
- Boitel, B. *et al.* *J. exp. Med.* **175**, 765–777 (1992).
- Chen, Z. W. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).

A way into the packaging

Hélène Richard-Foy

THE chromatin nucleosomal core particle presents a formidable obstacle to a regulatory protein trying to get at its target DNA for the purposes of activating transcription. The accessibility of this DNA is restricted by virtue of its intimate packaging with histone proteins in the nucleosome. Now two papers on pages 477 and 481 of this issue^{1,2} describe a human homologue of the yeast SWI/SNF complex which is instrumental in reorganizing nucleosome structure in the promoter region so that transcription factors can bind to their cognate sites on DNA.

This interaction of diffusible regulatory proteins with their DNA targets is probably achieved by local disruption of chromatin structure over short stretches of DNA. It has been proposed that, in yeast, a group of five functionally related nuclear proteins, SWI1/ADR6, SWI2/SNF2, SWI3, SNF5 and SNF6 (known as the SWI/SNF proteins), are involved in chromatin destabilization (see ref. 3 for review). SWI2/SNF2 belongs to a protein family that shares a predicted DNA-dependent ATPase domain⁴ and homologues have been cloned in higher eukaryotes such as *Drosophila* and humans (reviewed in ref. 5). The SWI/SNF proteins are global activators of a large set of inducible genes and are believed to form part of a multiprotein complex⁴, but how these nuclear proteins work, bearing in mind that they have only a weak intrinsic affinity for DNA⁶, is uncertain.

Genetic studies have indicated that SWI/SNF proteins may act by inducing changes in chromatin structure (reviewed in ref. 3). Analysis of the chromatin

architecture around the yeast *SUC2* gene promoter, which requires SNF proteins for its activation, showed that SWI2/SNF2 and SNF5 function by altering the structure of two nucleosomes, a rearrangement that is not a direct result of transcriptional activation⁷. Although this suggested that remodelling chromatin structure increases the accessibility of its DNA to diffusible regulatory proteins, it was not clear whether the SWI/SNF complex was acting directly on chromatin.

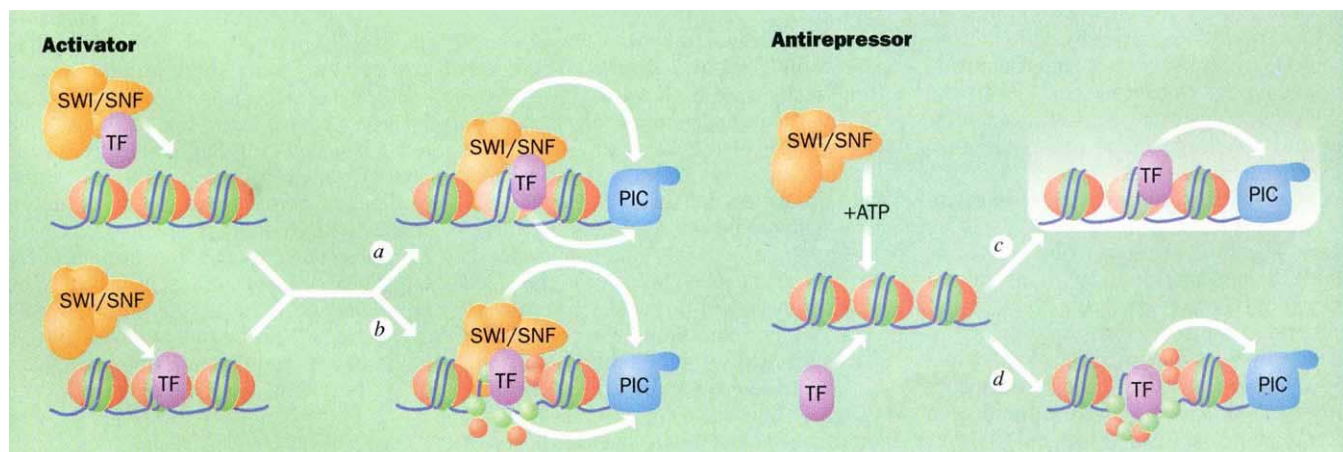
SNF proteins fused to heterologous DNA-binding domains from transcriptional activators can efficiently activate promoters containing their cognate DNA-binding sites^{8,9}, suggesting that once it has found its DNA target, the SWI/SNF complex behaves as a classic transcriptional activator. The idea thus arose that SWI/SNF DNA targeting could be achieved through its interaction with factors that are themselves sequence-specific regulators. This proposal was supported by the findings that the glucocorticoid receptor, a sequence-specific regulator, interacts with SWI3 *in vitro* and that SWI3 is necessary for promoter activation by the glucocorticoid receptor¹⁰. But from these experiments it looked more as though SWI/SNF was having a stimulatory effect on the assembly of the preinitiation complex rather than altering the structure. So if SWI/SNF proteins function not by interacting with a specific DNA sequence, but instead by assisting regulatory proteins that bind to DNA⁴, then what is the nature of this alliance?

The SWI/SNF multiprotein complex, which shows DNA-dependent ATPase

activity, has been isolated recently from yeast^{6,11} and now from humans^{1,2}, offering the opportunity for *in vitro* investigation and insight into how SWI/SNF acts. Kwon *et al.*¹ and Imbalzano *et al.*² have analysed the effect of the human homologue of the SWI/SNF complex (hSWI/SNF) on the structure of a positioned mononucleosome that contains a DNA-binding protein target sequence. The positioning of the core octamer histone on the DNA allowed the authors to study the influence of rotational positioning of the binding element on interaction with its cognate binding protein and they were able to show that hSWI/SNF induces an ATP-dependent structural change in the nucleosome. This is revealed by a change in the DNase I digestion profile: the pattern of the 10-base-pair periodic cut generated by the presence of a positioned nucleosome changed to a pattern resembling, but distinct from, that of naked DNA.

The important conclusion was reached that SWI/SNF remodels, but does not disrupt, the nucleosome structure. These two papers^{1,2}, and one reporting similar results for yeast⁶, together show that the SWI/SNF complex acts by hydrolysing ATP to promote dramatically the binding of specific or general transcriptional activators such as GAL4 or TATA-binding protein (TBP). This points to chromatin, and not the transcriptional machinery, as the primary target for the SWI/SNF complex. But it would not account for the limited number of genes that are involved in such regulation *in vivo*, because nucleosome remodelling was evident in the absence of promoter-specific and general transcription factors and showed no DNA-template specificity *in vitro*.

Another striking observation² was that,



Models for SWI/SNF transcriptional activation. SWI/SNF may act either as an activator or as an antirepressor. In the 'activator' model, SWI/SNF binds to chromatin by interacting with a specific DNA-binding activator, or transcription factor (TF). This coupling may take place before or after binding to DNA (left and right, respectively). In both cases binding of SWI/SNF will stimulate transcription (probably by acting on the preinitiation complex (PIC)). This may involve steps a or b, which rep-

resents a chromatin structure disruption. In the antirepressor scheme, SWI/SNF first disrupts the chromatin structure locally, the energy being provided by ATP hydrolysis. This chromatin structure rearrangement could be a remodelling (c) or a complete disruption (d) of the nucleosome. In both cases this will allow binding of the transcription factor TF and transcriptional activation. The highlighted sketch represents the most likely model.

upon SWI/SNF-induced chromatin remodelling, TBP was able to gain access only to promoters with a specifically positioned TATA sequence, so nucleosome positioning over regulatory sequences might be responsible for the specificity of the response. The association of the SWI/SNF complex with DNA-binding proteins that are sequence-specific regulators may also be implicated in the specificity of the process. Such DNA-binding proteins must be able to recognize their target at the surface of a nucleosome. The chromatin remodelling resulting from the interaction of this complex should then either enhance binding of this factor to its target

or, more likely, allow binding of another factor to a target that was originally hidden.

Investigations using the whole cell should now provide insights into the specificity of chromatin remodelling over regulated promoters. It may be that nucleosomes with different degrees of DNA accessibility positioned over such promoters will emerge. □

Hélène Richard-Foy is in the Laboratoire de Biologie Moléculaire des Eucaryotes du Centre National de la Recherche Scientifique, 118 route de Narbonne, 31062 Toulouse, France.

1. Kwon, H., Imbalzano, A. N., Khavari, P. A., Kingston, R. E. & Green, M. R. *Nature* **370**, 477–481 (1994).
2. Imbalzano, A. N., Kwon, H., Green, M. R. & Kingston, R. E. *Nature* **370**, 481–485 (1994).
3. Winston, F. & Carlson, M. *Trends Genet.* **8**, 387–391 (1992).
4. Peterson, C. L. & Herskowitz, I. *Cell* **68**, 573–583 (1992).
5. Carlson, M. & Laurent, B. C. *Curr. Opin. Cell Biol.* **6**, 396–402 (1994).

6. Côté, J. *et al. Science* **265**, 53–60 (1994).
7. Hirschhorn, J. N. *et al. Genes Dev.* **6**, 2288–2298 (1992).
8. Laurent, B. C., Treitel, M. A. & Carlson, M. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2687–2691 (1991).
9. Muchardt, C. & Yaniv, M. *EMBO J.* **12**, 4279–4290 (1993).
10. Yoshinaga, S. K. *et al. Science* **258**, 1598–1604 (1992).
11. Cairns, B. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **91**, 1950–1954 (1994).

CATALYSIS

How to sugar the pill

John M. Brown and Stephen G. Davies

NATURE has had billions of years to devise good production facilities (enzymes) for the synthesis of single enantiomers (mirror image forms) of certain molecules (asymmetric synthesis). Over the past two decades chemists have successfully addressed this same problem and have recently come up with many intriguing and exciting solutions. Where an asymmetric synthesis is required for molecules (for instance pharmaceuticals, agrochemicals or materials for electronics), one can be devised at least on the small scale. Now the challenge for the chemist is in making such processes economically and environmentally acceptable on a manufacturing scale. On page 449 of this issue¹, Wan and Davis demonstrate an elegant new approach to the problem which may turn out to be commercially viable.

Great strides have been made in asymmetric synthesis over the past two or three years. The ability to start from a molecule which is achiral (identical to its mirror image) and convert it into a single enantiomeric product (a molecule which is distinct from its mirror image, only one of the two being formed) is far advanced, although still a long way from the degree of perfection chemists would seek. To give a couple of examples: since Jacobsen's work² starting in 1990, it has been possible to convert many simple alkenes into a single enantiomer of their epoxides; the new oxygen atom is delivered by the manganese catalyst to a single face of the alkene double bond. And Corey has

developed an original discovery by Itsuno to provide a simple and specific method for the reduction of ketones, again giving a single enantiomer (for a review see ref. 3).

It happens that these simple functional groups are at the core of synthetic organic chemistry; epoxides and secondary alcohols are building blocks around which many syntheses can be devised. Add to these the earlier development of asymmetric hydrogenation⁴ and a few other asymmetric syntheses involving chemical catalysts, and you are *en route* to something fairly significant: a general purpose toolkit for the synthesis of enantiomerically pure molecules.

It sounds amazing, and to an extent it is; that is as far as the academic scientist or the drug discovery chemist is concerned. The implication is that quick and expedient single enantiomer routes to target molecules using purely chemical methods are now available. This avoids the need for a resolution (a physical separation of enantiomers, combining them with an enantiomerically pure 'handle' which then has to be removed) or chiral chromatography (limited to small scales if astronomical costs are to be avoided). But if the drug discovery chemist is successful, then it will be necessary to carry out this synthesis on a grander scale, and new challenges emerge. The problem for the process chemist is no longer how to conduct a perfect asymmetric synthesis, but how to do it easily and quickly on

a large scale, isolate the product simply, avoid contamination by the last traces of a metal catalyst which could percolate through to the final drug, and control the production of all waste materials. Some would say that this is the most difficult part.

Having the catalyst and reactants in different phases is an obvious way to proceed, permitting the desired separation by simple filtration processes. Unfortunately, heterogeneous asymmetric catalysis has a far poorer track record than homogeneous catalysis, with one or two limited successes in hydrogenation of reactants with a restricted range of functional groups. That situation may well change, but until it does, 'supported' homogeneous catalysts look more promising. It would be well to explain this apparent contradiction in terms. The catalyst consists of a reaction centre (the engine) connected to ligands which control the orientation of the reactant (the steering). These ligands can be adsorbed in a permeable solid material such as Montmorillonite or chemically bound to a polymer such as a macroreticular polystyrene. Although feasible, this has had a disappointing history. Too often the catalyst leaches slowly out of the support and defeats the object of the exercise; alternatively, it may be tightly bound but slow to react.

Another promising approach is to devise two-phase systems: make the catalyst water-soluble and the reactants insoluble and find a way of easily transporting the reactants across the interface. This technique has had considerable success in catalysis generally⁵, but not yet in asymmetric synthesis, where the results are no better than in organic media, and often worse.

The work by Wan and Davis¹ breaks new ground for heterogeneous analogues of homogeneous catalysts. The availability of glass micro-beads with controlled pore size has come about largely through solid-phase oligonucleotide synthesis, and such beads have near ideal characteristics for adsorbing a polar liquid (ethylene glycol) which together with the chiral ruthenium complex forms the catalyst phase. The reductant is dissolved in a nonpolar solvent and reaction requires interfacial transfer. Evidently this is no great barrier because the rate is about one-third of the homogeneous system in methanol, and the enantiomeric excess is 96 per cent (98:2 ratio of the two

1. Wan, K. T. & Davis, M. E. *Nature* **370**, 449–450 (1994).
2. Jacobsen, E. N. in *Catalytic Asymmetric Synthesis* (ed. Ojima, I.) 159–203 (VCH, Weinheim, 1993).
3. Wallbaum, S. & Martens, J. *Tetr. Asymm.* **3**, 1475–1504 (1992).
4. Takaya, H., Ohta, T. & Noyori, R. in *Catalytic Asymmetric Synthesis* (ed. Ojima, I.) 1–33 (VCH, Weinheim, 1993).
5. Herrmann, W. A. & Kohlpaintner, C. W. *Angew. Chem. int. Edn Engl.* **32**, 1524–1544 (1993).

enantiomers) in both cases.

(S)-naproxen is an important anti-inflammatory agent and much work has been devoted to developing asymmetric hydrogenation routes to this commercially significant compound. Wan and Davis have chosen wisely to demonstrate their new methodology in this well worked area where it compares favourably. The beauty of their method is that it uses a combination of the ideas of both homogeneous and heterogeneous asymmetric catalysis to create a useful technology.

This is just one example, and the generality remains to be established. But there is obvious promise for extension, perhaps most obviously to oxidation processes such as asymmetric epoxidations and dihydroxylations, and we may have seen the first example of a genuinely useful and practical new idea. □

John M. Brown and Stephen G. Davies are at the Dyson Perrins Laboratory, University of Oxford, South Parks Road, Oxford OX1 3QY, UK.

NEUROBIOLOGY

Catching the culprit prion

Konrad Beyreuther and Colin L. Masters

PRIONS, the infectious agents of certain neurological diseases, replicate like viruses. But they seem to be devoid of nucleic acid genomes and consist largely of PrP^{Sc} protein, a protease-resistant form of the normal cellular protein denoted PrP^C. The latest insight into the conversion of PrP^C into PrP^{Sc} is reported by Kocisko *et al.* on page 471 of this issue¹. What occurs naturally — the formation of protease-resistant prion protein — has now been achieved *in vitro* in a cell-free system composed of PrP^C and using PrP^{Sc} as 'seed'.

Prion diseases are rare degenerative neurological disorders that are inherited or sporadic in origin² and which can be transmitted³. They afflict humans (Creutzfeldt–Jakob disease, Gerstmann–Sträussler–Scheinker syndrome, kuru and fatal familial insomnia); sheep and goats (scrapie); cattle (bovine spongiform encephalopathy or BSE, otherwise known as 'mad cow disease'); mink (transmissible mink encephalopathy); mule deer, elk and antelope (chronic wasting disease); and domestic cats. Clearly, these diseases are of considerable medical and veterinary interest.

More than a decade ago the prion protein PrP^{Sc} was discovered in fractions enriched for scrapie infectivity⁴ and brain amyloid⁵. This led to the discovery that PrP^{Sc} is derived from the normal host protein PrP^C which is necessary for transmission and pathogenesis of disease⁶. In addition, in the case of the inherited human prion diseases, mutations identified in the *Prn-p* gene of patients were found to segregate with the disease⁷. This suggested that the mutations are sufficient to cause both the disease and the conformational conversion of PrP^C into PrP^{Sc}.

More information on the structural requirements for this conversion came from studies with transgenic mice, which showed that the conversion is most efficient when the amino-acid sequences of PrP^C and PrP^{Sc} are identical⁸. The trans-

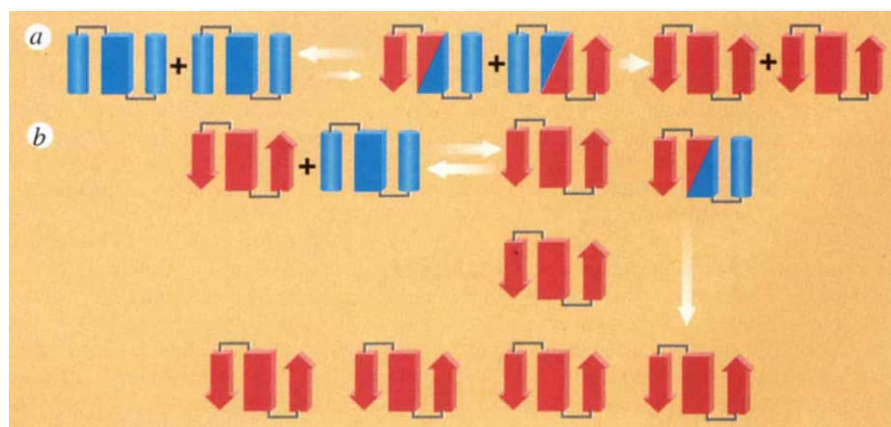
mission of prions from one species to another with non-identical sequences is possible, but it usually requires a prolonged incubation time for the initial trans-species passage. Subsequent passages are considerably shorter, however, suggesting that the sequences of PrP^C and PrP^{Sc} may become identical.

How then is PrP^{Sc} produced? Until now the evidence was that PrP^{Sc}, which is defined biochemically by its insolubility, resistance to protease and high β -sheet content, is derived post-translationally from the normal, protease-sensitive protein PrP^C. Generation of PrP^{Sc} seems to involve conformational rather than covalent modifications of the protein backbone^{9,10}. PrP^C has a high α -helical content and very low content of β -sheet, whereas PrP^{Sc} has a high β -sheet content

and is lower in α -helix than PrP^C (see figure)¹⁰. Because PrP^{Sc} had to be isolated from diseased brain or infected cells, the question of whether PrP^{Sc} formation requires oligomerization could not be addressed experimentally. The insolubility of PrP^{Sc} isolates did not allow the determination of its physical state, and also precluded analysis of the interaction of these isolates with host PrP^C in cell-free mixing experiments. Such analyses could be the first step towards an assay that might reveal whether a disease such as bovine spongiform encephalopathy does indeed constitute a health risk for humans.

The BSE outbreak, together with the iatrogenic transmissions of Creutzfeldt–Jakob disease in recipients of cadaver-derived growth hormone, fertility hormones or dura mater, has stimulated a hot debate about the transmissibility, prevalence and clinical variability of prion diseases. Public interest in BSE, a disease that was first recognized in Holstein Friesian dairy cattle in 1986, also reflects the financial implications that the disease has had for the British dairy industry. Inoculation of brain material from confirmed clinical cases of BSE has produced disease in sheep, goats, pigs, cattle, marmosets and mice, but not in hamsters¹¹. Whether human PrP^C can be converted to PrP^{Sc} by the BSE PrP^{Sc} agent¹² is not known.

That may change with the appearance of the paper by Kocisko *et al.*¹ and their assay for the cell-free formation of protease-resistant prion protein. The assay was established for radiolabelled hamster PrP^C purified by immunoprecipi-



Model for the conformational change of PrP^C to PrP^{Sc}. PrP^C has a high content of α -helix (42%) and probably no β -sheet (3%), and is shown in blue with two α -helical columns; by contrast PrP^{Sc} (red) has a high β -sheet content of 43% (shown as ribbon arrows) and an α -helix content of 30%. *a*, The spontaneous conversion of PrP^C to PrP^{Sc} is very slow, and only one in a million normal individuals develops sporadic Creutzfeldt–Jakob disease. However, familial PrP mutations increase the likelihood of this conversion, and all carriers of mutations who live beyond the age of onset develop the disease. *b*, As Kocisko *et al.*¹ show, conversion of PrP^C to the protease-resistant form is fast in the presence of PrP^{Sc} as a 'seed'. It can be observed within two days in the cell-free system, whereas degeneration of the central nervous system may take at least 125 days to develop in transgenic mice⁸. Because formation of protease-resistant prion protein requires the addition of PrP^{Sc} that has some native structure, the direct PrP^C–PrP^{Sc} interaction probably involves β -strand stabilization in PrP^C and intermolecular hydrogen-bonding between β -strands of PrP^{Sc} and PrP^C (ribbons).

tation from cells transfected with a retroviral expression system. This PrP^c was then incubated with highly purified but unlabelled PrP^{Sc} from scrapie-infected hamster brain. Incubation of [³⁵S]-labelled recombinant hamster PrP^c with PrP^{Sc} for two days, and subsequent treatment with proteinase K, produced the characteristic ³⁵S-labelled protease-resistant PrP fragments which were similar in size to those obtained for PrP^{Sc} from scrapie isolates.

Interestingly, partial denaturation of PrP^{Sc} by pretreatment in 3 M guanidine-HCl gave higher conversion efficiencies but appeared not to be mandatory. Kocisko *et al.* also show that the conversion does not require the biosynthesis of new PrP^c, or its amino-linked glycosylation, or the presence of its normal glycosphosphatidylinositol anchor. This implies that direct protein-protein interaction, and possibly oligomer formation between PrP^{Sc} and PrP^c, is sufficient for the conversion reaction to occur.

These observations raise several questions. What is the relationship between PrP^{Sc} formation and prion replication, and does the conformational change of PrP^c to PrP^{Sc} also produce infectivity? Can this conversion reaction be accelerated by exogenous catalysts such as metal ions¹³?

The limitation of the present version of the cell-free conversion assay is the requirement of at least a 50-fold excess of PrP^{Sc} over PrP^c. This high excess of infectivity introduced with the conformational catalyst PrP^{Sc} precludes the detection of any newly generated scrapie infectivity (prions) which may be associated with the radiolabelled protease-resistant forms. If one of the components can be used in immobilized form, or if the amount of the PrP^{Sc} required can be reduced by other means, then what looks like a specialized and limited experimental model system could become a key to the detection of infectious prions and to understanding the pathogenesis of prion diseases. □

Konrad Beyreuther is in the Laboratory for Molecular Neuropathology, ZMBH, University of Heidelberg, Im Neuenheimer Feld 282, D-69120 Heidelberg, Germany. Colin L. Masters is in the Department of Pathology, University of Melbourne, Parkville, Victoria 3052, Australia.

Dirty clouds and global cooling

Graeme L. Stephens

THE increasing amount of sulphate aerosol gathering in the atmosphere, by-product of the anthropogenic increase of greenhouse gases, is likely to have considerable effects on our climate. That was the conclusion reached by Taylor and Penner¹ earlier this year, when they considered the complex cooling patterns that might be produced as aerosol particles reflect solar radiation into space (see

–0.3 W m⁻². Calculations are mainly limited by the assumptions that must be made for how much anthropogenic aerosol exists in the atmosphere (above normal background amounts) and where it is most concentrated around the globe.

Still more telling, if Jones *et al.* are right, is the indirect effect. The idea stems from that of Twomey *et al.*⁷: when clouds form in air with increased aerosol concentrations, the cloud reflectivity is enhanced because of the increased concentrations of cloud condensation nuclei (these are the tiny centres on which droplets grow). Polluted clouds are thought to be composed of larger concentrations of smaller droplets than clean clouds, and it is essentially the increased concentration of droplets in these clouds that makes them more efficient at reflecting solar radiation. It is difficult to get at the magnitude of the Twomey effect, and it has been demonstrated only in limited ways (by looking at the effects of effluents from ship tracks on cloud albedo)⁸. But recent estimates by Boucher *et al.*⁹ and here by Jones *et al.* suggest that the climate forcing due to the indirect aerosol effect exceeds that of the direct effect. Jones *et al.*, using a climate model that includes a relationship between aerosol concentrations and cloud droplet concentrations, estimate this forcing to be about –1.3 W m⁻², although they are quick to highlight the uncertainty of this value. Both Boucher *et al.* and Jones *et al.* arrive at their estimates using climate models that are run with and without the parameterized effects of sulphate aerosol on clouds.

The story of the impact of anthropogenic aerosols on climate impact does not end there. A further indirect effect stems from their influence on cloud microphysics. Because the droplet distributions are shifted towards smaller sizes, polluted clouds are less likely to produce drizzle and thus are less likely to rain out. Prolonging these low-level clouds by shutting off the drizzle formation leads to an increase in cloud cover and thus an increase in the reflection from these regions of cloud¹⁰. Although calculating the contribution of this subtle effect is difficult, global estimates have been made by Boucher *et al.*⁹ based on the use of a climate model that includes cloud microphysical processes; and their preliminary results suggest that it is roughly as large as the Twomey effect.

When all effects of anthropogenically produced aerosol are taken together, then the predicted global forcing is comparable in magnitude to the direct forcing by greenhouse gases. It is a grave mistake in logic, however, to expect that these aero-

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Sensitive subjects — oceanic clouds are vulnerable to aerosol pollution.

News and Views by Wigley²). But the effects are twofold, and on page 450 of this issue Jones *et al.*³ make a brave attempt to assess the subtler role of aerosols — their influence on the microphysical properties of clouds.

First let's take the effects of greenhouse gases. It is widely believed that increasing concentrations of these radiatively active gases will lead to global warming as less infrared radiation escapes to space and a greater flux strikes the Earth's surface. The change in the state of the fluxes at the top of the atmosphere is referred to as the radiative forcing of climate, and current estimates⁴ suggest that the increase in carbon dioxide since the dawn of the industrial age has produced a forcing of about 2–2.5 W m⁻².

The effect of sulphate aerosols on the Earth's radiative budget is complicated. Sulphate particles directly influence the radiative balance of the planet by increasing the reflection of solar radiation to space. Estimates of the resulting globally averaged forcing vary^{1,5,6} from –0.9 to

Images Colour Library

- Kocisko, D. A. *et al.* *Nature* **370**, 471–474 (1994).
- Masters, C. L. *et al.* *Brain* **104**, 535–558 (1981).
- Gibbs, C. J. *et al.* *Science* **161**, 388–389 (1968).
- Bolton, D. C., McKinley, M. P. & Prusiner, S. B. *Science* **218**, 1309–1311 (1982).
- Prusiner, S. B. *et al.* *Cell* **35**, 349–358 (1983).
- Bueller, H. *et al.* *Cell* **73**, 1339–1347 (1993).
- Hsiao, K. *et al.* *Nature* **338**, 342–345 (1989).
- Scott, M. *et al.* *Cell* **73**, 979–988 (1993).
- Hope, J. *et al.* *EMBO J.* **5**, 2591–2597 (1986).
- Pan, K.-M. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **90**, 10962–10966 (1993).
- Bradley, R. *Dev. Biol. Stand.* **80**, 157–170 (1993).
- Hope, J. *et al.* *Nature* **336**, 390–397 (1988).
- Bush, A. I. *et al.* *Science* (in the press).

sol effects simply compensate the greenhouse warming tendency. As Taylor and Penner¹ described earlier this year, one of the main reasons that a simple cancellation of effects is unlikely is because aerosol is regionally concentrated. Estimating the regional character of the direct forcing is relatively straightforward: the largest forcing occurs near the most prolific sources of anthropogenic aerosols. Determining the regional response even to this forcing is another matter, as Taylor and Penner found from their general circulation model results.

The regional pattern of the indirect forcing is much more complicated than that of direct forcing. Jones *et al.* tend to reinforce Twomey's¹¹ analyses, suggesting that the clouds most noticeably affected by pollution will be those found in clean marine environments such as over the Southern Oceans; in areas already fairly polluted, the effects will level out. This result illustrates how the immediate climate forcing by these sulphate aerosols may be far removed from local sources. Overall, contrasting the radiative and microphysical properties of clean Southern Ocean clouds with their Northern Hemispheric counterparts should help test our understanding of the indirect aerosol forcing.

Essential to the results of both Boucher *et al.*⁹ and Jones *et al.* is the knowledge gleaned from detailed measurements of cloud and aerosol microphysics obtained from aircraft flying in and around clouds. For example, the results of Martin and Johnson¹² obtained from several flights by the UK Meteorological Office's C130 research aircraft gave the concentrations of droplets and condensation nuclei used in these studies. Aircraft observational programmes have a valuable contribution to make to our understanding of global climate. □

Graeme L. Stephens is in the Department of Atmospheric Sciences at Colorado State University, Fort Collins, Colorado 80523, USA.

Cradle at the ribosome

Walter Neupert and Roland Lill

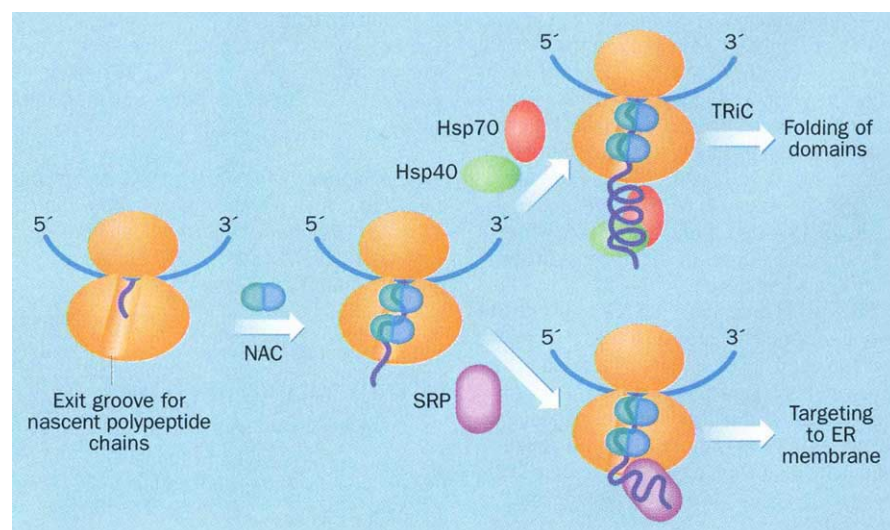
It is widely thought that, at the beginning of protein synthesis, nascent polypeptide chains on ribosomes are sequestered in a 'tunnel' or 'channel' where they are shielded from the environment^{1,2}. Interaction of nascent chains with components that determine their further fate is believed to occur after more than 40 amino-acid residues have been strung together. Molecular chaperones then bind to nascent polypeptides as they start to fold³; alternatively, targeting factors are recruited that direct proteins to the various organelles of the cell.

On page 434 of this issue⁴, Wiedmann *et al.* report the identification and isolation of a protein complex that interacts with nascent polypeptide chains on ribosomes in the eukaryotic cytosol at a very early stage — that is, with segments of nascent chains that are about 30 or even fewer amino-acid residues from the peptidyl-transferase. In a crosslinking approach, using a photoactivatable crosslinker incorporated in the nascent chains, they detected a protein of relative molecular mass 33K which became covalently linked to a variety of nascent chains, provided that these chains were associated with the ribosomes. The 33K protein is not a core component of the ribosomes (it could be removed by salt treatment). Binding to such 'stripped' ribosomes containing nascent chains served as an assay for the purification, from bovine brain cytosol, of a protein complex that contained not only the 33K protein, but also a 21K protein. Wiedmann *et al.* call the heterodimeric

complex NAC (for nascent-polypeptide-associated complex), and it consists of α -NAC and β -NAC. NAC is relatively abundant, being present at micromolar concentrations in the cytosol of various cells analysed.

To determine more precisely when and where NAC interacts with the ribosome and nascent chain, the authors compared crosslinking of NAC to nascent cytosolic proteins with crosslinking to nascent secretory proteins carrying signal sequences at their amino terminus. Where NAC was present in the translation system, both kinds of nascent chains were crosslinked to NAC; when sufficiently long, the signal-bearing chains were also crosslinked to the 54K component of the signal recognition particle (SRP)⁵⁻⁷. In previous studies, SRP54 has been found to interact with nascent chains longer than about 70 amino acids.

Surprisingly, after depletion of NAC, added SRP became crosslinked not only to secretory nascent chains but also to non-secretory nascent chains; in addition, SRP could be crosslinked to chains that were much shorter than 70 residues. When NAC was added back to the depleted translation system it re-bound to the ribosome-associated nascent chains and prevented crosslinking of SRP to nascent chains lacking signal sequences. Even more surprisingly, after binding to SRP in the absence of NAC, non-signal-bearing nascent chains could not only be targeted to microsomes but also became translocated across the membrane. This



Components proposed to interact with nascent polypeptide chains. The first known interaction of segments of a nascent polypeptide chain takes place with NAC, a heterodimer composed of α - and β -subunits. After the growing polypeptide chain emerges from the ribosome, chaperones Hsp40 and Hsp70 can bind to initiate folding of protein domains in conjunction with other chaperones like TRiC. When an endoplasmic reticulum (ER) signal sequence is released from NAC, SRP can bind to it and serve to deliver the ribosome-SRP complex to the ER membrane.

1. Taylor, K. E. & Penner, J. E. *Nature* **369**, 734–737 (1994).
2. Wigley, T. M. L. *Nature* **369**, 709–710 (1994).
3. Jones, A., Roberts, D. L. & Slingo, A. *Nature* **370**, 450–453 (1994).
4. *IPCC Climate Change: The IPCC Scientific Assessment* (Cambridge Univ. Press, 1990).
5. Charlson, R. J., Langner, J., Rodhe, H., Leovy, C. B. & Warren, S. G. *Tellus* **43**, 1252 (1991).
6. Kiehl, J. T. & Briegleb, B. P. *Science* **260**, 311–314 (1993).
7. Twomey, S. A., Piepgrass, M. & Wolfe, T. L. *Tellus* **B36**, 356–366 (1984).
8. Coakley, J. A. Jr, Bernstein, R. L. & Durkee, P. A. *Science* **237**, 1020 (1987).
9. Boucher, O., Le Treut, H. & Baker, M. B. *Proc. Eighth AMS Conf. Atmos. Radiation*, Nashville, Tennessee, 558–560 (Am. Meteor. Soc., Massachusetts, 1994).
10. Albrecht, B. A. *Science* **245**, 1227–1230 (1989).
11. Twomey, S. A. *Atmos. Env.* **A25**, 2435–2442 (1991).
12. Martin, G. M. & Johnson, D. W. *Proc. 11th int. Conf. Clouds and Precipitation*, Montreal **1**, 158–161 (Am. Meteor. Soc., Massachusetts, 1992).

latter process was not very efficient, but the authors present a number of criteria to suggest that it was specific.

According to Wiedmann *et al.*, NAC may have a general function of binding to nascent polypeptides as they emerge from the peptidyltransferase centre. It may shield the nascent chains against improper folding or interaction with other components, and may help form the supposed 'channel' or 'tunnel' in the large ribosomal subunit. The new findings suggest that this tunnel is more like a groove with a removable cover, in part or totally composed of NAC molecules (see figure). NAC does not bind to more distal segments of a nascent chain as it needs the ribosomal context for binding to occur. The SRP would normally bind to signal sequences as soon as they are released from NAC. In the artificial situation of NAC depletion, SRP would interact nonspecifically with non-signal sequences and thereby facilitate targeting and eventual translocation of non-secretory proteins into the endoplasmic reticulum (ER).

The implications of these conclusions are far-reaching, both with regard to folding of newly synthesized chains and to the mechanism of protein translocation into the ER.

NAC may be the first or one of the first components that a growing polypeptide chain encounters before being delivered to components of the protein-folding machinery, such as the molecular chaperones Hsp70 and Hsp40 (see figure). More chaperones such as TRiC may come into play soon thereafter⁸. Clearly, the proposed function of NAC has to be supported by *in vivo* studies using appropriate mutants. It will have to be determined how NAC interacts with both the ribosome and the nascent chains, what the binding partners on the ribosome are, and to which part of a polypeptide chain NAC is binding. NAC-like molecules might be expected to exist with ribosomes other than those present in the eukaryotic cytosol, such as those in chloroplasts, in mitochondria and — not least — in bacteria. If NAC lives up to the role proposed for it, this will be a major step forward in the understanding of protein synthesis, folding and sorting.

The observations that SRP can bind to non-signal peptides and that the peptides are targeted to the ER will raise some controversy. Taking them at face value, they would mean that binding of SRP to signal sequences is more a default reaction than the result of a specific recognition process. NAC would have a large part in determining SRP specificity. Even more provocatively, the data could be interpreted to suggest the participation of the signal sequence in targeting to the ER membrane only, and not in translocation. The decision of whether a protein is destined for translocation into the ER or to remain

in the cytosol would then be made not at the membrane but in the cytosol.

That a polypeptide chain could initiate translocation without a signal seems heretical at present. On the other hand, it is fair to say that the role of the signal in initiating ER translocation is all but understood. In fact, Wiedmann *et al.* can cite in their favour a report that signal sequences are not required for secretion of polypeptides in cells of *Escherichia coli* harbouring a mutant SecY protein in their pre-protein translocase⁹. Because individual steps in targeting and translocation may be subject to some lack of specificity owing to the degenerate nature of signal sequences, quantitative measurements, and certainly experiments with intact cells, will be necessary to make a convincing case that signals have at most an auxiliary function in ER translocation.

The report by Wiedmann *et al.* raises other questions. Can proteins fold into their native conformation if NAC is missing during translation? How are secretory proteins targeted which follow a post-translational pathway and which have been shown not to require SRP *in vivo* or *in vitro*¹⁰? Does NAC also help in ensuring specificity of sorting of pre-proteins to other membrane systems, such as chloroplasts and mitochondria?

Furthermore, one might speculate whether equivalents of NAC exist in these organelles where the passage of unfolded polypeptides through the membrane in many respects resembles the passage through the ribosomal groove. Would such organellar NAC equivalents keep incoming, unfolded segments of pre-proteins in such a conformation that they could then interact with chaperones and with sorting factors that target the incoming polypeptide to the organelle's subcompartments? It will be interesting to see whether a whole set of hitherto unknown factors are required for what is currently thought to be a spontaneous process. □

Walter Neupert and Roland Lill are in the Institute for Physiological Chemistry, Physical Biochemistry and Cell Biology, University of Munich, Goethestrasse 33, D-80336 Munich, Germany.

1. Yonath, A. *Rev. Biophys. biomol. Struct.* **21**, 77–93 (1992).
2. Johnson, A. E. *Trends biochem. Sci.* **18**, 456–458 (1993).
3. Ellis, R. J. *Nature* **370**, 96–97 (1994).
4. Wiedmann, B., Sakai, H., Davis, T. A. & Wiedmann, M. *Nature* **370**, 434–440 (1994).
5. Wolin, S. L. *Cell* **77**, 787–790 (1994).
6. Krieg, U., Walter, P. & Johnson, A. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8604–8608 (1986).
7. Kurzchalia, T. V. *et al. Nature* **320**, 634–636 (1986).
8. Frydman, J., Nimmesgern, E., Ohtsuka, K. & Hartl, F. U. *Nature* **370**, 111–117 (1994).
9. Derman, A. I., Puziss, J. W., Bassford, P. J. & Beckwith, J. *EMBO J.* **12**, 879–888 (1993).
10. Klappa, P., Müller, G., Schlenstedt, G., Wiech, H. & Zimmermann, R. in *Membrane Biogenesis and Protein Targeting* (eds Neupert, W. & Lill, R.) 137–146 (Elsevier, Amsterdam, 1992).

Sex or violence

PRESIDENT Clinton, like so many male politicians these days, faces repeated hostile probes into his sexual life. This seems unfair. In all primates, including man, the whole biological point of fighting your way up to being top male is to have your pick of females. As Daedalus discussed last week, sex and social dominance are both mediated by one hormone. The ambitious politician, urged on by testosterone but cramped by the new puritanism, is thus trapped in Wood's double bind: 'if you've got the qualifications, you'd be mad to apply for the job'.

So Daedalus wants to uncouple the two drives fuelled by testosterone. The hormone must bind to two different sets of receptors in the brain: one for sex, the other for power and dominance. Receptors with such different functions can hardly be identical. So 'tweaked' variants of testosterone might be designed to bind preferentially to one or the other. In this connection Daedalus recalls the theory that a smelly molecule stimulates its receptor by means of its molecular vibrations. This mechanism may well apply to internal biochemical receptors. If so, then testosterone could be tweaked, not by risky chemical changes, but simply by isotopic substitution. It would be chemically unaltered, but its changed atomic masses would shift its vibrational modes.

Most mammals synthesize their sex hormones from cholesterol. DREADCO's chemists are now feeding test animals on various isotopically labelled cholesterols; these will be converted *in vivo* to labelled testosterone. Some should stimulate the creatures' fighting instincts; others will be peaceably aphrodisiac. When he has found the most efficient labellings, Daedalus will copy them in tablets for human use; ideally not cholesterol itself (health lobbyists would complain) but something further downstream in the biosynthesis.

The cruel double bind of the new puritanism will relax at last. DREADCO's 'Pure Power' tablets will channel the user's testosterone entirely into social aggression. Delighting in the meanest political in-fighting and ruthless wheeler-dealing, he will be unmoved by the most beguiling bimbos. Policemen and soldiers in mixed-sex forces, required to be lethally aggressive while treating their female colleagues with the most delicate courtesy, will welcome them too. By contrast, DREADCO's 'Peace'n'Love' tablets will appeal to nostalgic hippies. Their testosterone will all go into sexual enterprise, leaving them wallowing ecstatically at the bottom of the social ladder.

David Jones

Legibility in lecture theatres

SIR — There is a striking discrepancy between the legibility of ordinary printed text and that of information commonly presented to the most-distant viewer in a lecture theatre. Here I propose a new standard for writing size that will greatly improve the readability of such material.

The usual criterion for legibility is that the height of a displayed letter should subtend an angle not less than some minimum value δ at the eye of the most-distant viewer. The potential amount of legible information that can be presented is therefore proportional to the solid angle that the display subtends at the eye of the most-distant viewer. It is convenient to measure this capability in terms of the information index $I = A/X^2$, where A is the presentation area and X is the distance to the most-distant viewer¹.

For a constant amount of information to be communicated at a given standard of legibility, the height of the presentation area must increase in proportion to X . For overhead projection screens the usual criterion has been that this height should not be less than 1/6 of X (ref. 2). This fraction is known as the screen design factor. Normally, the image of a projection table just fills the screen, so that I is equal to the square of the screen design factor. Values of I for the usual range of screen design factor are shown in the table. An information index of 70 millisteradian (msr) seems to be adequate for a general-purpose lecture theatre, consis-

INFORMATION INDEX OF OVERHEAD PROJECTION SCREENS

Screen design factor	1/5	1/5.3	1/6	1/7	1/8
Information index(msr)	40	36	28	20	16

tent with dual overhead projection screens of 1/5.3 screen design factor.

For printed text, publishers usually ensure that the height of lettering intended for normal reading does not fall below 1.5 mm. Thus at the normal 250-mm reading distance, $\delta = 6$ milliradian (mrad). For footnotes the height is allowed to fall to 1 mm ($\delta \geq 4$ mrad). The current legibility standard for projected images of lettering^{3,4} is $\delta \geq 2.5$ mrad, and for the symbols used on optometric eye charts, the Snellen criterion for 6-6 vision is that the threshold of legibility is 1.5 mrad. So for a large lecture theatre of $X = 15$ m, the height of lettering for the most-distant viewer must be 90 mm for normal reading standard, 60 mm for footnote standard, 38 mm for the projection criterion and 23 mm for threshold of legibility.

Chalkboard writing in large theatres is often 30–40-mm high, which to the most distant viewer will appear smaller than footnotes of ordinary printed text.

In comparison, consider an overhead projection system with a 3-m screen and a 285-mm projection table, in the same 15-m-deep theatre. The system magnification is 10.5, and with a transparency lettering height of 7 mm, the height of the projected image will be 74 mm. For the most-distant viewer this value is much closer to the normal reading standard and with proper control of ambient lighting^{3,5} the contrast will be higher.

Consider now a transparency having lettering of height S that subtends an angle δ at the eye when viewed directly at a distance R (the direct reading distance). If the transparency is placed on an overhead projection table of dimension T and imaged on a screen of design factor $1/n$ (so that the image of the projection table just fills the screen), then $S/\delta = Tn = R$. This is the fundamental equation relating all legibility parameters for overhead projection systems¹.

The significance of R is that transparencies viewed directly at this distance are as legible as the projected image will be to the most-distant viewer in a standard lecture theatre. Legibility can thus be checked during preparation

of a transparency.

The recommendation for a screen design factor of 1/6 originated at a time when $T = 254$ mm ($R = 1.524$ m). For modern projectors of $T = 285$ mm, a screen design factor of $n = 5.3$ is necessary for the same legibility. It has been proposed therefore that 1/5.3 should be adopted as the standard for the screen design factor in lecture theatres¹. Because excellent inexpensive projectors are now widely available, I propose that a new legibility standard of $\delta \geq 5$ mrad should be adopted, which is closer to the normal reading standard, and that the original $\delta \geq 2.5$ mrad standard should be reserved as an absolute minimum. The approximate lettering heights on overhead transparencies would then be 7 mm for normal use and 4 mm as the absolute minimum.

E. W. Dearden

Department of Physics,
University of Queensland,
Brisbane 4072, Australia

1. Dearden, E. W. J. *archit. Plan. Res.* **7**, 21–32 (1990).
2. Taylor, J. *The Science Lecture Room* (Cambridge Univ. Press, 1967).
3. TerLouw, A. L. *Legibility Standards for Projected Material* (Eastman Kodak, New York, 1965).
4. *Legibility: Artwork to Screen* (Eastman Kodak, New York, 1983).
5. *IES Lighting Guide: Lecture Theatres* (Illuminating Engineering Society, London, 1973).

Interstellar detection of C_{60}^+

SIR — The report of the possible detection of C_{60}^+ in the interstellar medium¹ is exciting, but brings to light the difficulties of unambiguous identification. The only spectroscopic information at present available on ions of relevance to the identification of the carriers of diffuse interstellar bands² are absorption measurements in neon matrices³. An exception are the known optical spectra of some 200 polyatomic cations in the gas phase, but these are mainly of conjugated polynes and halo-substituted benzenes⁴.

Although it is well known that the matrix observations are subject to a solvation shift with respect to the isolated state, and that the linewidths are not relevant for a gas-phase comparison, four important criteria, if fulfilled, make an identification more plausible. We have come to formulate these in conjunction with our studies of the highly unsaturated hydrocarbon radicals⁵ for just these reasons: present availability of matrix but not gas-phase data; and for purposes of astrophysical predictions.

(1) In matrix absorption spectra, short vibrational progressions in the excited state are often seen. Thus, the separation of the origin band and the vibrational satellite in the excited state should be within 5–10 cm^{-1} of the separation in the gas phase of the related diffuse interstellar bands. By this means the gas-matrix shift

is reduced to this uncertainty.

(2) The ratio of the relative intensities of the two absorption peaks in the matrix should be similar to the ratio of the equivalent widths for the diffuse interstellar bands (except for forbidden transition, where anomalous intensities may be induced in the matrix).

(3) The half-widths (FWHM) of the diffuse interstellar bands themselves should be practically the same if the pair of bands belong to the same carrier. This test is independent of the matrix data.

(4) The gas-matrix shift should be consistent with known experimental information.

Not all these criteria have been addressed or answered in ref. 1. Specifically, the FWHM of the two features appear not to have been sufficiently well determined, to decide if they are comparable. The ratio of the equivalent widths is similar to the intensity ratio of the matrix absorptions, but here one has to be aware that the two corresponding peaks in the matrix spectrum vary in ratio over the range 1.5–2 to 1, depending on the production conditions⁶, and were attributed to two distorted forms of C_{60}^+ .

In the matrix adsorption spectrum of C_{60}^+ , a peak corresponding to vibrational excitation in the upper electronic state is apparent, 235 cm^{-1} to higher energy of the origin band and about one-sixth of its

intensity⁶. Therefore, one should try to detect a diffuse interstellar band at $9,419 \pm 9 \text{ \AA}$. Its equivalent width should scale to the reported band at $9,632 \text{ \AA}$ as in the matrix spectrum and should possess a comparable FWHM.

Finally, the apparent gas-matrix shift is smaller than expected. This sort of information has only recently become available for the lowest electronic transitions of C_{60} and C_{70} , from a comparison of gas-phase⁷ and neon-matrix⁸ spectra, and is about 50cm^{-1} for both. Consequently, a comparable if not larger shift is expected for their cations. If the reported diffuse interstellar bands at $9,632$ and $9,577 \text{ \AA}$ are of C_{60}^+ , then the shift is only 14 or 7cm^{-1} , respectively.

John P. Maier

*Institute of Physical Chemistry,
University of Basel,
Klingelbergstrasse 80,
4056 Basel, Switzerland*

1. Foing, B. H. & Ehrenfreund, P. *Nature* **369**, 296 (1994).
2. Herbig, G. *Astrophys J.* **196**, 129–160 (1975).
3. Salama, F. & Allamandola, L. *Nature* **358**, 42–43 (1992).
4. Maier, J. P. *Chem. Soc. Rev.* **17**, 45–67 (1988).
5. Fulara, J. *et al.* *Nature* **366**, 439–441 (1993).
6. Fulara, J. *et al.* *Chem. Phys. Lett.* **211**, 227–234 (1993).
7. Haufner, R. E. *et al.* *J. chem. Phys.* **95**, 2197 (1991).
8. Fulara, J. *et al.* *Chem. Phys. Lett.* **206**, 203–209 (1993).

The early life of the European eel

SIR — The early life history of the European eel *Anguilla anguilla* is still something of a mystery¹. After growing for a long time (from 5 to 12–15 years) in freshwater habitats, yellow eels metamorphose into silver eels, then as they mature they migrate in order to breed in the western Atlantic. As early as 1922 (ref. 2), the Danish oceanographer Johannes Schmidt outlined the spawning area of both Atlantic eels *A. anguilla* and the American eel *A. rostrata* in the Sargasso Sea ($23\text{--}30^\circ \text{N}$; $48\text{--}74^\circ \text{W}$); the larvae, called leptocephali, migrate towards the European continental shelf, where they metamorphose into glass eels which enter brackish or fresh waters. The duration of this transatlantic migration was estimated to be 2–3 years². Here we show that in fact this migration lasts for less than one year by reading the otoliths, a sensory device that acts as a sort of 'black box' in the fish, recording information from hatching to death through rhythms of calcium mineralization.

Experimental rearing of leptocephali to study their early growth has not been successful: larvae die within 5 days after induced breeding of *A. anguilla* or *A. japonica*. The discovery of daily increments within the otolith (sagittae) microstructure³ presents an alternative method to follow the different stages of the early life of teleost species.

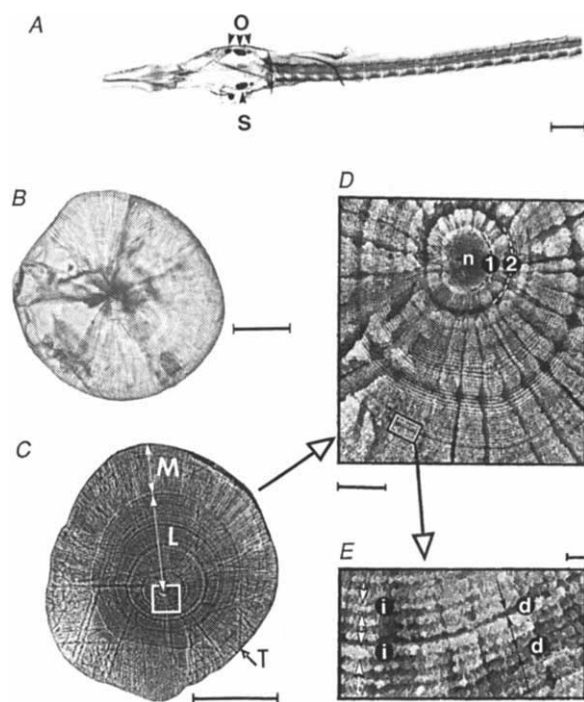
We investigated otolith microstructure at a later stage of larval growth, in glass eels, to obtain a record of the previous larval life history and hence to find out how long European leptocephali take to cross the North Atlantic from the spawning area in the Sargasso Sea to the European continental shelf, and their age when they enter estuarine waters as glass eels.

We used scanning electron microscopy to examine 1,000 sagittae of glass eels caught along the coast from Morocco to the Netherlands, and from western Mediterranean coasts. Growth rate was estimated on the assumption that growth increments were deposited daily, which we verified in the case of the Japanese eel.

All sagittae displayed clear daily increments (see figure) distributed concentrically around the core. This pattern was common to the other eel species *A. rostrata*, *A. japonica* and *A. marmorata*. The mean age of glass eels ranged from 190 to 280 days. The calculated growth rate was $0.26\text{--}0.30$ mm per day. Thus, European eel larvae spend less than one year in transatlantic migration.

This result disagrees with Schmidt's estimate of 2–3 years and suggests that leptocephali larvae do not simply drift passively with the Gulf Stream. Oceanographic data give the velocity of Atlantic currents supposedly used by leptocephali during their voyage towards Europe: water masses are estimated to be translated from the western to the eastern Atlantic over a period of almost one year⁴. The duration of the larval migration (up to 280 days) is therefore shorter than that of the transfer of the water masses, so leptocephali must be capable of swimming actively.

This larval transatlantic migration of European eels is a surprising biological event. Other eel species have spawning areas close to the continental shelves, so their larval migration is fairly short, 4–5 months for *A. japonica* and *A. marmorata*. Can the long and complex larval migration⁵ of the European eel be explained? Improvements in our understanding of the formation of the North Atlantic have thrown new light on the biology of these remarkable fish, several



A, Microradiograph of the glass-eel's head showing the three pairs of otoliths. Sagittae are the biggest ones. Scale bar, 1 mm. **B**, Sagitta of glass eel, *in toto*, at an optical level. Scale bar, 100 μm . **C**, Scanning electron micrograph illustrating internal microstructure of a glass-eel sagitta. The ground and etched otolith is circular in shape, displaying a concentric ring pattern with the core region (first layer) delimited by an etched groove. A second layer ($80\text{--}100 \mu\text{m}$) comprised a series of growth increments corresponds to the leptocephalus growth zone, L. A third layer ($50\text{--}70 \mu\text{m}$) beginning with diffuse rings and ending with a typical double ring (the transition, T) corresponds to the metamorphosis zone, M. Scale bar, 100 μm . **D**, Detail of the first layer. The nucleus, n, is visible as: 1, a deep cavity (length $11.7 \pm 2.5 \mu\text{m}$, width $9.8 \pm 1.8 \mu\text{m}$) surrounded by a dark ring, the end of yolk sac resorption; 2, a crystalline crown ($20.1 \pm 1.7 \mu\text{m}$, $17.8 \pm 2.1 \mu\text{m}$) which is in turn surrounded by a deep dark groove, mouth functional, first ingestion of food. Scale bar, 10 μm . **E**, A daily increment is composed of a bipartite structure: incremental zone i, white ring (white arrowheads); discontinuous zone d, dark groove (black arrows). Scale bar, 1 μm .

authors having suggested a relationship between eel migration and Wegner's continental drift theory. The North Atlantic was still relatively narrow during the Tertiary period and the spawning areas of eel species were presumably situated in the same geographical location as the present Sargasso Sea. As the two continents (America and Europe) surrounding the North Atlantic began to drift apart, so the eels had to undertake journeys of steadily increasing length in order to reach their spawning area.

Raymonde Lecomte-Finiger

*Laboratoire de Biologie Marine,
Université de Perpignan,
52 Avenue de Villeneuve,
66860 Perpignan Cedex, France*

1. Kleckner, R. C. & McCleave, J. J. *mar. Res.* **46**, 647–667 (1988).
2. Schmidt, J. *Phil. Trans. R. Soc.* **B211**, 179–208 (1922).
3. Pannella, G. *Science* **173**, 1124–1127 (1971).
4. Lecomte-Finiger, R. *Mar. Biol.* **114**, 205–210 (1992).
5. McCleave, J. J. *Fish Biol.* **43** (suppl. A), 243–273 (1993).

Reaching for the stars

Donald D. Clayton

Home is Where the Wind Blows: Chapters from a Cosmologist's Life. By Fred Hoyle. University Science Books: 1994. Pp. 443. \$32.50, £18.95.

FRED Hoyle has for five decades questioned entrenched dogma and replaced it with reasoned alternatives. What good fortune to have this beautifully written autobiography of one of this century's leading scientists. Astrophysics graduates after 1972 have had little opportunity to experience his aura, to see how much he created or to perceive how much he influenced the field. They may not understand what prompted George Gamow's famous limerick ending: "And God said 'let there be Hoyle', and there was light". Reading Hoyle's engaging book will not fill the gap, however, because it is written with great modesty and humorous self-deprecation. This is a distinctly human biography of an exceedingly talented man; and because he has resisted the temptation to vindicate his career choices, giving no more space to his theory of nucleosynthesis in stars, for example, than to the problems of growing up in the world as we find it, his recollections are exceedingly graceful. Every astronomer will want to read this work, and those who do will be well rewarded.

Hoyle straddled the research worlds of the new physics, initially as a Cambridge student of the 1930s, and of astronomy after the Second World War. Without despair he notes: "my generation [of theoretical physicists] was the unluckiest of the past half century — too late to receive anything but crumbs from the rich table around 1926, too early for quantum electrodynamics and quarks, and six years [half a generation for theoretical physics] lost to wartime". What Hoyle envisaged was the application of the new physics to a greatly expanded view of the Universe. He achieved this on many fronts, opening up our thinking about cosmology as relativistic particles and fields, in contrast to simply geometrical universes. The post-war era was the appropriate moment for this to happen, after the formulation of quantum mechanics and general relativity, but before radio-source counting, the discovery of the 3 °K microwave background and the theory of nucleosynthesis in stars. In the historical sense, Hoyle's ideas were bound to come, but it was by no means determined that they would take root from the impoverished West Riding of Yorkshire in a man of singular sensitivity, wit and humanity, but also a stubborn man who raged throughout his life against stupid artificial rules and oppressions.

The autobiography is laid out in three sections: "34 Primrose Lane", in which

Hoyle recounts his childhood and university education; "The Larger World of Science", in which he recounts his work on radar during the Second World War, his return to Cambridge and his invention of nucleosynthesis; and "Home is Where the Wind Blows", a glimpse into the mind of the mature cosmologist having to deal



Moving in different circles — Fred Hoyle at Stonehenge, England

with the politics of science.

The first part, written in the mid-1980s, is a reprinting (with photographs this time) of Hoyle's delightful earlier account of childhood, *The Small World of Fred Hoyle* (reviewed in *New Scientist* by Geoffrey Burbidge on 28 August 1986). From a literary standpoint, this is the most memorable writing, though not of the most direct scientific interest. It is beautifully done, and one would not change a single word. Hoyle describes with gentle humour and deep insight his path from a bright bad schoolboy to a graduate with a Cambridge first. The increasing unacceptability of his teachers' methods and behaviour forced him to drop out of school at the age of eight, a conscientious objection supported strongly by his mother. From a second school he won a scholarship to Bingley Grammar School, near Bradford in West Yorkshire, centre of coal mines and woollen mills. Avoiding pitfalls and Catch-22 situations, he eventually won a scholarship to Emmanuel College, Cambridge, further discovering and proving his great creative gifts until finally reaching the top, winning the Mayhew Prize in theoretical physics of the Mathematical Tripos of 1936 and then, in 1938, the Smith Prize. Hoyle earned a fellowship at St John's College, where he did research with Rudolph Peierls (until Peierls moved to Birmingham in 1937) and then Paul Dirac, "because [Dirac] could not resist the circular counterlogic

of a supervisor who didn't want a research student who didn't want a supervisor". Hoyle provides much interesting history of the University of Cambridge in the 1930s, exposing even the seeds of alienation that would culminate in his exasperation in 1972 with the political demagogues of that institution, itself so clearly loved by Hoyle. At several reveries of Part One I found myself moved by his emotional insights, especially into the transitory nature of personal relationships, which often seem so eternal only to fade into haunting memories. On one level, this part depicts the life of a smart northern English lad between the shadows of the wars, a period piece and a regional

portrait. On another level it describes the struggle of a headstrong talent, despising pompous authoritarianism, finding his own voice and the depth of his creativity, and proving himself the equal of any student in Cambridge. Here also, recounted with touching humour, is the story of how he finds and wins his wife, Barbara Clark, who provides the one great lasting relationship of his life.

The second part begins with a sympathetic and detailed portrait of Arthur Eddington, whose chair, the Plumian professorship, Hoyle would come to fill. Astronomers will value Hoyle's view of Eddington's clever evasion of uncertainty over the role of nuclear power in stars by using not only the mass-luminosity relation for stars but their surface temperatures as well, and how Hoyle and R. A. Lyttleton were able to improve on this by actually calculating the surface temperature using the hydrogen-burning reactions identified by Hans Bethe. Hoyle also reminds readers of the escape from the 'iron sun' blind alley. Mindful of the eventual fiasco of Cambridge politics, Hoyle speculates: "Eddington's death was of much greater consequence for me than I realized at the time. If events had been allowed to develop naturally in postwar years, I think my relations with the British astronomical establishment might well have been greatly changed."

Like many scientists, Hoyle's priorities were forcibly changed by the Second World War, and two chapters relate

From Home is Where the Wind Blows

directly to the early years of the conflict. Hoyle worked on radar at Nutbourne, the large field in Sussex where the radar antennas for naval radar were designed. He describes the problems of returning in 1945 to Cambridge with a junior lectureship and of defining his direction in research. But define it he did with characteristic creativity. His account of his theory of nucleosynthesis in stars is required reading for astronomers. Coming well before his steady-state cosmology, it shows clearly that we should bury once and for all the canard that stellar nucleosynthesis was presented to prop up the steady-state theory. One has heard that too often, and it is false.

In the third part, Sir Fred shares much of the administrative side of his scientific life. Elected Plumian Professor in 1958, he became the champion of the revitalization of British astronomy, its greatest fundraiser, political liaison for government planning in astronomy, founder of an instantaneously famous research institute, chairman for the design and construction of the Anglo-Australian Telescope, and perhaps the most spell-binding lecturer astronomy has known. He, of course, does not claim these superlatives; but as a half-time colleague at his institute in Cambridge during its magic six years (1966–72), I can assure the reader that they are no exaggeration. These pages lead us through the high points and dwell little on the lows. Of special interest are Hoyle's efforts to create and maintain the Institute of Theoretical Astronomy on Madingley Road. Never able to suffer fools gladly, Hoyle succumbed in 1972 to destructive acts by envious university competitors. A hard pull up a Scottish Munro was his technique for clearing his mind and keeping it in balance, as I was lucky to learn from him. His love of the hills is evident throughout — even in the title.

In the last chapter, Hoyle gives us in forceful prose not only the definitive history of the steady-state Universe but also an evolution for its future. He explains in direct language why he regards simple Big Bang cosmology as a dead end (and this despite his lasting fame for naming the Big Bang and for his pioneering calculation of the origin of helium in the event). His scalar field interprets black holes as sources of matter rather than as sinks. Like a mountain guide with a compass, he outlines the cosmological terrain over which he has travelled for half a century, complete with its wrong turns and acrimonious assaults by others. It is an inspired vision, with testable scientific history far more distant than that single unknowable event of creation just one Hubble time ago. □

Donald D. Clayton is in the Department of Physics and Astronomy, Clemson University, Clemson, South Carolina 29634, USA.

A misanthropic philanthropist

John L. Heilbron

Alfred Nobel: A Biography. By Kenne Fant. Arcade: 1993. Pp. 342. \$24.95.

ALFRED Nobel flaunted inconsistencies. The King of Dynamite, the inventor of smokeless powder and the mastermind of the Swedish munitions industry, he supported peace movements; an engineer, applied chemist and aggressive capitalist, he wrote, and established a prize for writing, "literature of an idealist tendency". With a good head for business but a weak stomach for delegation, he ran his

scientist and clever entrepreneur.

Fant does not try to integrate the many Nobels at whom he hints with the "grumpy" (such was Nobel's self-characterization in his letters to Sofie) man immanent in the documents he quotes. From the opening quotation, from a letter of 1887 ("Alfred Nobel — pitiful creature, ought to have been suffocated at birth by a humane physician when he made his howling entrance into this life"); through epigraphs from "Nemesis" ("My life is woven out of torment"); to excerpts from the letters to Sofie ("My stomach hurts so much"; "My nobility of soul has withered away"; "I hope that you will never be struck by the feeling of debasement that embitters my days"), Fant's Nobel is permanently lugubrious and self-absorbed. This stasis may be related to Fant's profession of cinematography. The establishment of character in a film requires or rewards techniques different from those appropriate to a written biography.

Readers of *Nature* may be disappointed by Fant's perfunctory accounts of Nobel's inventions and their place in the second Industrial Revolution. He does not explain how the chemists of the mid-nineteenth century understood the properties of nitroglycerine or describe how dynamite was manufactured, nor does he tell in any detail how Nobel's explosives revolutioned civil engineering and improved the art of war.

The letters of Sofie Hess have only recently been made available. They have survived in a safe place because of blackmail: Nobel's executor, Ragnar Sohlman, bought them from her during the contest over Nobel's will. Sohlman reasoned that the news that Nobel had kept a silly, spendthrift, ignorant girl (so Nobel describes Sofie in his letters to her) would create a public image inimical to the implementation of the odd idealistic testament of a man then already sufficiently suspect as a merchant of death. Fant's book demonstrates Sohlman's wisdom. The letters, printed as they are without commentary, create a picture of Nobel that swamps the more interesting and complicated personality implicit in other documents.

Fant's good partial portrait demonstrates the desirability of a full biography of the King of Dynamite that would put together the elements of Nobel's personality and integrate the whole with his knack for invention and approach to business. Such a biography would make a welcome appearance some time between the centenary of Nobel's death (1996) and that of the first Nobel prizes (2001). □

J. L. Heilbron is in the Office for History of Science and Technology, University of California, Berkeley, California 94720, USA.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Alfred Nobel (1833–96).

enterprises by correspondence without an office or a secretary. Nobel insisted that he was frail, but worked for 18 hours a day; he went into society, and acted the misanthrope; he prided himself on his rationality, and took elaborate precautions against being buried alive.

A good writer with an ear for quotation can scarcely avoid making Nobel's life interesting. Kenne Fant is such a writer. He lets Nobel speak primarily through letters to his brothers and to Sofie Hess, a woman 25 years his junior whom he kept for more than a decade. Fant also relies heavily on a play Nobel wrote a few years before his death. Its title, "Nemesis", is the most cheerful thing about it. The man who emerges from this selection of documents is a melancholic, hypochondriacal, misanthropic workaholic. To be sure, Fant lets slips hints of other Nobels: the good son, brother and uncle; the philanthropist; the man of wit, friend of Victor Hugo and lover of literature; the creative

The story of e

Jerry P. King

e : The Story of a Number. By Eli Maor. Princeton University Press: 1994. Pp. 223. \$24.95, £19.95.

MATHEMATICIANS tell this tale: a man questions the cost of his life insurance. An actuary tells him that insurance costs are based on life expectancy, which is measured with probability distributions, particularly the normal distribution.

"What is a normal distribution?" the man asks.

"It's this," the actuary says, showing the man a graph of a bell-shaped curve.

"What are those symbols?"

"That's the equation which describes the curve."

"OK, but what is that?" asks the man pointing.

"That's the number pi."

"What the devil is pi?"

"Pi is the ratio of the circumference of any circle to its diameter," says the actuary.

"That's crazy," the man says. "What can circles have to do with how long I will live?"

It's a nice story and if you tell it properly you will get a chuckle. But the tale has always seemed to me contrived. To be sure, π appears in the equation of the normal distribution. But so does the number e . In fact, e appears more prominently than π because the normal distribution is only the graph of a suitably modified exponential function. Moreover, e wanders ubiquitously throughout the mathematical world. Its various roles include the exponential function e^x and the base for the natural logarithm $\ln(x)$. Often π and e show together in the same expression as in the famous Euler identity: $e^{i\pi} + 1 = 0$, which links in a single breathtaking equation the five important constants e , i , π , 1 and 0 . Like π , e is irrational, which means it has no finite or infinitely repeating decimal expansion. Also like π , e is transcendental, which means that it is not the root of any polynomial with integer coefficients. Given all this, why did the man in the story ask only about π ? Why did he not ask about e ?

The answer — as the story indicates — is that π can be explained in terms of circles while no analogous, simple explanation exists for e . A real person might well ask the actuary about e . But when you tell the story you have the man point only to π . To explain e you need to write a book. Until now, that is.

Eli Maor set himself a difficult task. He says: "My goal is to tell the story of e on a level accessible to readers with only a modest background in mathematics". Whether he has succeeded depends on

one's interpretation of the word modest. I think he falls slightly short.

The trouble is that e is an analytic notion. It belongs to the branch of mathematics called analysis and is defined directly in terms of the basic calculus concept of limit. To deal with e , one must deal with the notion of limit: Maor does this pretty well and begins with a financial interpretation of e : invest \$1 at 100 per cent yearly interest. Compound the interest once yearly and the account contains \$2 at the year's end; compound it twice and it contains \$2.25. If the interest is compounded 10^3 , 10^4 , 10^5 or 10^6 times, the year-end dollar balance becomes, respectively, 2.71692, 2.71815, 2.71827 and 2.71828. Continuing this process gives a sequence of numbers that seems to have a limit of 2.71828, which is the correct value of e to five decimal places.

But mathematics knows not seems, and Maor properly wants to present more

rigorously the ideas that surround e . Although he relegates some sophisticated notions to appendices, the body of the book goes as far as calculus, differential equations and complex analysis. His presentation is not self-contained, so the reader's background is critical.

On the other hand Maor wonderfully tells the story of e . The chronological history allows excursions into the lives of people involved with the development of this fascinating number. Maor hangs his story on a string of people stretching from Archimedes to David Hilbert. And by presenting mathematics in terms of the humans who produced it, he places the subject where it belongs — squarely in the centre of the humanities. □

Jerry P. King is in the Department of Mathematics, Lehigh University, 14 East Packer Avenue, Bethlehem, Pennsylvania 18015-3174, USA.

The Sun and its variability

David W. Hughes

Discovering the Secrets of the Sun. By Rudolf Kippenhahn. Wiley: 1994. Pp. 262. £49.95, \$79.95 (hbk); £19.95, \$31.95 (pbk).

Solar and Stellar Activity Cycles. By Peter R. Wilson. Cambridge University Press: 1994. Pp. 274. £40, \$59.95.

As stars go, the Sun has two enormous advantages for scientists. It is close, so its surface can be viewed in great detail, spatially, temporally and spectroscopically. It also varies — not so much that Earth would be uninhabitable, but certainly enough to make our neighbouring star extremely interesting. My old Oxford professor divided astronomers into those who investigated the Sun and those who looked at other things. I do not have to tell you which group he regarded as being the most important.

Rudolf Kippenhahn has written a first-class introduction to solar science. Even though it contains no equations, the author still confronts the reader with all the modern excitement of the subject. I like especially the chapters that concentrate on solar oscillations, the solar neutrino problem and the observation of the Sun from space. The author also introduces many topics from a historical perspective; the text reads like an adventure story as more and more clues are discovered with time, and the true picture of the Sun is slowly revealed.

The author is not afraid of magnetism. Sunspots, prominences and flares all illustrate the influence of strong magnetic fields on hot solar plasmas. The complications of the sunspot cycle and the reversals

of the solar magnetic fields underline the importance of those field lines that become twisted and frozen, and of the magnetized regions that become buoyant. The way in which the solar spin rate varies with both latitude and depth is also stressed.

Kippenhahn discusses the relationship between the Sun and Earth and considers the greenhouse effect and the rather tenuous link between Earth's climate and the phase of the sunspot cycle. He does, however, skate rather speedily over the huge computational effort that went into calculating the physical and chemical state of the solar interior. And little is made of the past origin of the Sun and its future evolution. He also gives the reader too little information on how the Sun compares in age, size and composition to the myriad other stars in the Galaxy. But these must be taken as minor criticisms. *Discovering the Secrets of the Sun* is the best introduction to solar science that I have read for decades and I shall insist that my students read it.

Even though the Sun is relatively easy to investigate, it still has many mysteries. Peter R. Wilson stresses these in *Solar and Stellar Activity Cycles* and he goes so far as to write: "it is not possible to provide a definitive account of the mechanisms underlying [solar] cyclic activity at the present time; opinions differ strongly on some aspects, whereas a general bafflement prevails in other areas".

It is fairly typical of the perversity of the Sun that, just as soon as one thinks one understands why sunspot activity varies with a 22-year periodicity, one then has to explain why spots disappeared altogether

between AD1642 and 1705. And one cannot even appeal to symmetry because the activity in the solar northern hemisphere is slightly different from that in the southern, and at high latitudes solar cycles overlap and it is difficult to tell whether certain features belong to the old cycle or the new one.

Wilson has provided the university student and postgraduate with a perfect mixture of what is known and what is still to be found out. There are detailed discussions about topics such as coronal holes, polar reversals, large-scale fields, non-linear dynamos, helioseismology, isorotational surfaces, the question of whether the sunspot cycle is chaotic, activity predictions and the general aperiodicity of solar activity. Not only is the book produced, illustrated and referenced to the normal high quality of the Cambridge Astrophysics Series, it is also extremely readable. □

David W. Hughes is in the Department of Physics, University of Sheffield, Sheffield S3 7RH, UK.

Nature's chemicals

Leslie Crombie

Scent and Fragrances. By Günther Ohloff. Springer: 1994. Pp. 238. DM128, £50.50, \$88.

Natural Products: Their Chemistry and Biological Significance. By J. Mann, R. S. Davidson, J. B. Hobbs, D. V. Banthorpe and J. B. Harborne. Longman: 1994. Pp. 455. £24.99, \$44.95 (pbk).

PLEASANT odours have been one of mankind's indulgences since first recorded time. Ancient Assyrian tablets record that incense was offered to the God of the Sun in Nineveh and that during the reign of Hammurahi a thousand talents (29,000 kg) were burnt at the Bel-temple of Babylon annually. Throughout the ages, expense has not been spared, and this is still true today. Bulgarian rose oil requires 800 man-hours of labour to pick the 3 tonnes of early morning flowers needed to produce 1 kg of the oil, which sells for around \$10,000. It takes eight million jasmine flowers weighing 1,000 kg to produce 1 kg of jasmine absolute oil, selling for upwards of \$15,000.

Of course, this is the exotic and expensive end of the perfumery industry. Application of organic chemistry has allowed identification of the essential odorous components of these plant mixtures and made them available more cheaply in synthetic form. Other cheap raw materials such as petrochemicals and the renewable resources of α - and β -pinene are today turned into the large quantities of fragrance materials, many of non-natural ori-

gin, required by modern society in its household articles such as soap and washing powder.

Scent and Fragrances is a delight to read and I can highly recommend it to chemists with interests in the perfumery industry. Although the large number of chemical formulae may seem at first

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Hebrew perfumes – cinnamon, balsam, nard and valerian (twelfth-century Persian manuscript).

sight forbidding, there is much in the book to interest nonchemists also: it is authoritative and written from first-hand experience. Apart from being a much respected organic chemist, Günther Ohloff has long experience of the industry and has contributed substantially to the study of the relationships between chemical structure and odour.

The early part of the book deals with chemoreceptors and structure–odour relationships. There is a certain elusiveness about the latter as the human population is far from uniform in its responses. A study involving six odour types and one-and-a-half million people has revealed some interesting facts. About 1.2 per cent of the population has no sense of smell (total anosmia). Females have a greater odour sensitivity than males, but this can vary during the menstrual cycle, while smoking can intensify certain odours and diminish others. Not surprisingly, odour perception diminishes above the age of 70, and drastically above 80. About a hundred compounds are known whose odours are not perceived by part of the population, that is, these people are partially odour-blind.

A large part of the book is concerned with odorants from natural sources and the information is concisely yet attractively presented, though in a few places an

update is needed because of rapid research progress (such as in the field of the jasmonoids). The author points out the fascinating relationships of named odorants to well-known perfume brands.

In one way or another, many organic chemists—including those working in the perfumery industry—first become seriously involved with natural products during their postgraduate careers and it would be interesting to know why most standard organic chemistry texts for undergraduates contain so little on this topic. Possibly the vast extent of the subject deters authors. So many specific chemical structures are available as possible illustrative examples that choice becomes bewildering. To control such choice, *Natural Products* is constructed around seven broad themes: carbohydrates, nucleic acids and their derivatives, amino acids and peptides, fatty acids, terpenoids, phenolics and alkaloids. The themes are handled in different ways by different hands and there is little attempt to coordinate the content or to deal with topics that do not fall into the chosen sections. The book seems best viewed as a compilation of seven monographs under one cover; the continuity and uniformity of judgement displayed by a good single-author work is lacking.

Nevertheless, these are all interesting and useful monographs. The choice of themes covers areas of current interdisciplinary interest such as carbohydrates and nucleic-acid derivatives, and I was impressed by the clear account of amino acids and peptides. There are perhaps missed opportunities in the section on fatty acids, where more could have been written on the biosynthetic importance of linoleic and linolenic hydroperoxides in plants, reducing the coverage of some of the well-trodden chemistry of the prostaglandins. The contribution on terpenes reflects the richness of this area for the illustration of organic reactions and their mechanisms, while biosynthetic themes, among others, are dealt with in the accounts of phenolics and alkaloids.

According to the cover, "*Natural Products* is written for final-year undergraduate and postgraduate students of organic and medicinal chemistry, biochemistry, pharmacy and pharmacology, as well as the professional market". This is a large constituency with differing scientific backgrounds, needs and aspirations, so the content of the book must be something of a compromise. It will be valuable to lecturers and some postgraduates as an information source, but it will seriously overload undergraduates, who need a much less sophisticated introduction to the subject of natural products. □

Leslie Crombie is in the Department of Chemistry, University of Nottingham, Nottingham NG7 2RD, UK

Chemical depletion of ozone in the Arctic lower stratosphere during winter 1992–93

G. L. Manney*, L. Froidevaux*, J. W. Waters*, R. W. Zurek*, W. G. Read*, L. S. Elson*, J. B. Kumer†, J. L. Mergenthaler†, A. E. Roche†, A. O'Neill‡, R. S. Harwood§, I. MacKenzie§ & R. Swinbank||

* Jet Propulsion Laboratory/California Institute of Technology, Pasadena, California 91109, USA

† Lockheed Palo Alto Research Laboratory, Palo Alto, California 94304, USA

‡ Centre for Global Atmospheric Modelling, Department of Meteorology, University of Reading, Reading RG6 2AU, UK

§ Department of Meteorology, University of Edinburgh, Edinburgh EH9 3JZ, UK

|| Meteorological Office, London Road, Bracknell RG12 2SZ, UK

Satellite observations of ozone and chlorine monoxide concentrations during winter 1992–1993 show that in February 1993 chlorine in the lower stratosphere was mostly in chemically reactive forms. Decreases in stratospheric ozone concentration during February and early March 1993 are consistent with chemical destruction by this reactive chlorine. Comparison with changes in the distribution of long-lived chemical and dynamical tracers shows that the observed decrease cannot have been caused solely by dynamical processes.

THE large ozone loss over Antarctica in late winter¹ is known to be caused by chlorine chemistry^{2–4}. Lower-stratospheric chlorine is converted to reactive forms by heterogeneous chemistry on polar stratospheric clouds (PSCs) followed by the action of sunlight^{2,5,6}. Enhanced abundances of chlorine monoxide (ClO), the dominant reactive form of chlorine in the daytime stratosphere, were observed during aircraft flights into the Arctic polar vortex in the 1989 and 1992 winters^{7–9}. Global measurements by the Upper Atmosphere Research Satellite (UARS) Microwave Limb Sounder (MLS) showed ClO abundances in the early January 1992 Arctic lower stratosphere comparable to those in the Antarctic during August¹⁰. Direct observation of Arctic ozone loss is more difficult than in the Antarctic due to greater dynamical variability in the northern polar vortex^{11–13}. The stronger Antarctic vortex isolates the air in which ozone destruction occurs, and is much colder than the Arctic vortex^{12–14}. Antarctic lower-stratospheric temperatures typically remain low enough for PSC formation for ~4 months, until after the spring equinox—as opposed to a few days to ~2 months in the Arctic¹⁵, where temperatures are well above the PSC threshold by the spring equinox. Chemical ozone losses in the Arctic should thus be smaller than in the Antarctic, and will be more difficult to detect against the climatological increase in polar ozone during Arctic winter.

To separate dynamical and chemical effects, we examine observations of ozone and ClO on isentropic surfaces (surfaces of constant potential temperature, θ) in relation to long-lived tracers, and to vortex structure as diagnosed from potential vorticity (PV)¹⁶. MLS observations of ozone and ClO during the 1992–93 Arctic winter, and their relation to the polar vortex, are compared with those in the 1991–92 Arctic and 1992 Antarctic winters, and with Nimbus 7 Limb Infrared Monitor of the Stratosphere (LIMS) ozone during the 1978–79 Arctic winter. Both the LIMS^{17,18} and the MLS^{19,20} data used here have been extensively examined and are sufficiently reliable for our examination of seasonal trends. The passive tracers nitrous oxide (N₂O) and methane (CH₄), which have no significant stratospheric sources and have chemical lifetimes long compared to dynamical timescales, were measured by the UARS Cryogenic Limb Array Etalon Spectrometer (CLAES)²¹ during the 1991–92 and 1992–93 winters. Our examination of the CLAES data used here indicates that they are also reliable for examining temporal trends

in and around the polar vortex. Both MLS and CLAES switch between viewing high northern and high southern latitudes every ~36 days.

Analyses of aircraft^{22–24} and MLS¹⁰ observations showed decreasing ozone in the Arctic lower stratosphere (~15–19 km) in January 1992; comparison here of MLS-observed ozone changes with expected transport effects supports the view that lower-stratospheric ozone in January 1992 was depleted by chlorine chemistry.

MLS measured enhanced ClO in the 1992–93 Arctic polar vortex beginning in early December 1992; by mid-February 1993, ClO mixing ratios >1 part per billion by volume (p.p.b.v.) were seen throughout the polar region¹⁰ and persisted until late February. Vortex-averaged ozone on the 465 K isentropic surface (referred to as 'ozone at 465 K' from here on) was observed to decrease by ~20% between mid-February and mid-March. Analyses of long-lived tracer data and vortex evolution show that this decrease could not have been caused solely by transport processes at this time. These results imply significant depletion of ozone by chlorine chemistry in the Arctic lower stratospheric vortex between mid-February and mid-March 1993.

Ozone in the Arctic vortex

MLS viewed high northern latitudes in December to early January, and mid-February to mid-March in 1991–92 and 1992–93. Figure 1 shows synoptic maps of ozone at 465 K (~19 km) in mid-February and mid-March for three Arctic winters (1979 from LIMS, and 1992 and 1993 from MLS), along with contours of Rossby–Ertel potential vorticity (PV) in the region of strong PV gradients. This region is coincident with strong westerly winds and demarcates the polar vortex. During 1979 and 1993, the vortex was considerably eroded by strong warmings between mid-February and mid-March. There were no strong warmings in late winter 1992, and the size of the 1992 vortex barely changed. Ozone at 465 K in 1992 and 1993 was largely confined within the vortex, in contrast to 1979, when PV gradients were much weaker and the vortex therefore less isolated throughout the winter. In fact, tongues of low ozone within the 1979 vortex in February suggest the possibility of intrusion of lower-latitude air into the vortex²⁵. Ozone mixing ratios peak in the mid-stratosphere, near ~35 km at high latitudes; thus, downward transport across isentropic surfaces, expected from diabatic cooling in the

winter stratosphere, should increase vortex ozone at 465 K. Ozone in the vortex increased during late winter in 1979 and 1992, as expected in the absence of significant chemical ozone loss. In 1993, however, ozone in the vortex decreased between mid-February and mid-March. Horizontal transport could

decrease vortex ozone because below ~ 550 K (~ 21 km), there is less ozone at low latitudes than in the polar vortex. However, Fig. 2 shows that the decrease in 1993 vortex ozone was fairly uniform in time, with no obvious signs (within the resolution of the data) of comparatively ozone-poor low-latitude air entering the vortex. Some ozone-rich air was observed being drawn off the vortex edge (for example, on 18 February and 2, 6 and 10 March), which could have decreased the area of high ozone in the vortex, and thus could alter a vortex average.

Vortex-averaged ozone, ClO and vortex minimum temperatures at 465 K for these three years are shown in Fig. 3. Overall, minimum temperatures were lowest in 1992–93 and highest in 1978–79. Although Arctic lower-stratospheric temperatures generally drop below the PSC formation threshold in only a small region of the vortex, in early January 1992¹⁰ and in February 1993 minimum temperatures are located away from the vortex centre, in a region experiencing both sunlight and strong winds; much of the vortex air thus moves through the region of cold temperatures, increasing the extent of PSC processing within the vortex. Unlike in the Antarctic, stratospheric warmings regularly raise temperatures in the Arctic vortex above the PSC formation threshold^{11,13}. This occurred with the onset of a strong warming in mid-January 1979, followed by a major warming in mid-

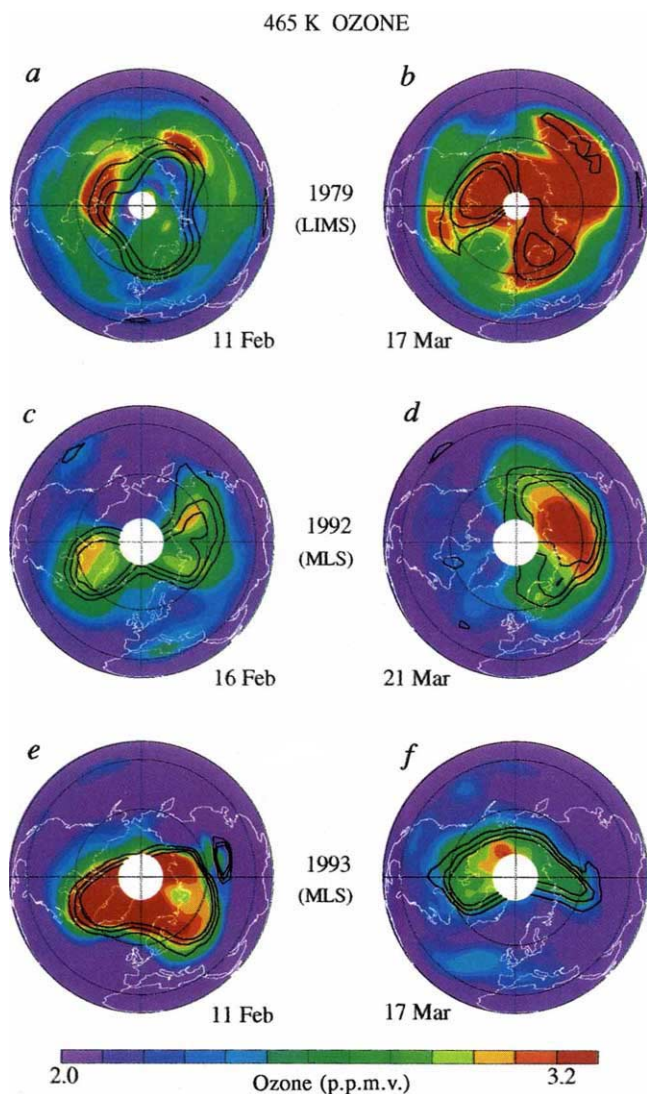


FIG. 1 *a–f*, Ozone on the 465 K isentropic surface for selected days in mid-February and mid-March 1979, 1992 and 1993. MLS data (*c–f*) have horizontal resolution of ~ 400 km and vertical resolution of ~ 4 km; the resolution of the LIMS data (*a, b*) is similar. Precisions (r.m.s.) of individual MLS ozone measurements for this data version (no. 3) are ~ 0.2 p.p.m.v., with absolute accuracies of ~ 15 – 20% . MLS data are gridded for 12:00 GMT using Fourier transform techniques that separate time and longitudinal variations³⁸. Potential vorticity (PV) contours of 2.5 , 3.0 and $3.5 \times 10^{-5} \text{ K m}^2 \text{ kg}^{-1} \text{ s}^{-1}$ are overlaid in black. PV is calculated¹³ from United Kingdom Meteorological Office (UKMO) assimilation system³⁹ temperatures and winds for 1992 and 1993, and from LIMS geopotential heights (using a nonlinear primitive equation balance to obtain winds⁴⁰) and temperatures for 1979. Comparison of PV calculated from US National Meteorological Center (NMC) data available from 1978 to the present⁴¹ indicates that the PV fields from different data sets used here are comparable; in particular, the relative strength of the PV gradients is not an artefact of the use of different datasets⁴². UKMO temperatures are used to interpolate gridded MLS data from pressure to isentropic surfaces (MLS temperatures are not currently available at these altitudes). The projection is orthographic, extending to the equator; thin dashed circles show 30° N and 60° N latitudes, and the Greenwich meridian is at the bottom of the plots.

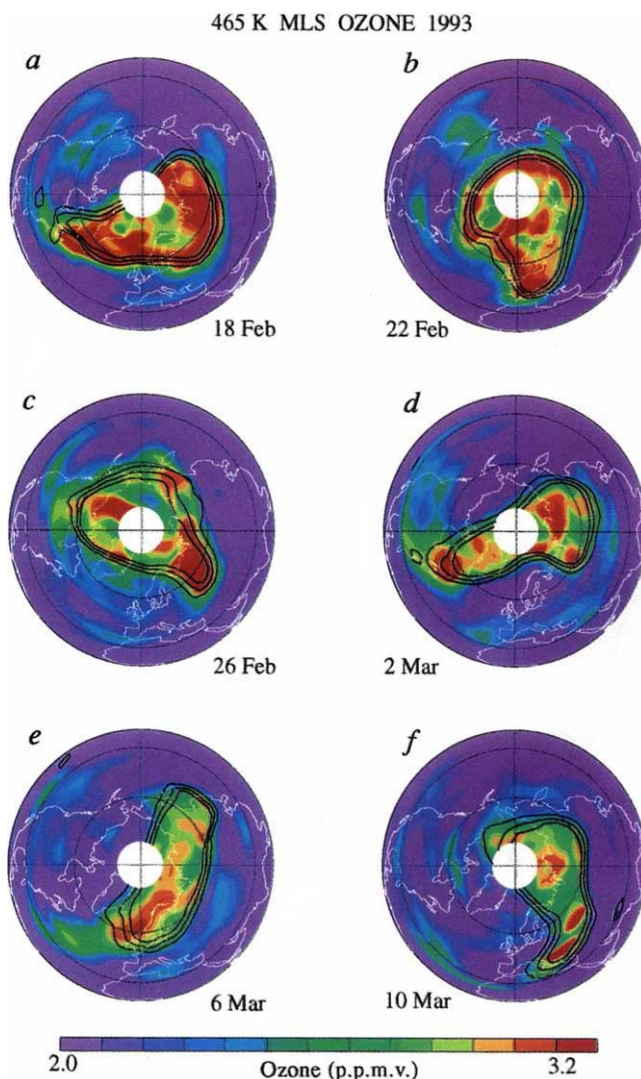


FIG. 2 *a–f*, MLS ozone on the 465 K isentropic surface at 4-day intervals from 18 February to 10 March 1993. Potential vorticity contours and layout are as in Fig. 1.

February¹⁸, and with the onset of a strong warming in mid-January 1992^{13,26}. The 1993 vortex, however, was relatively strong and cold through mid-February, with strong stratospheric warmings in late February and early March leading to the final warming²⁷. The persistence of low temperatures through late February 1993, a period of increasing sunlight at high northern latitudes, implies that high ClO concentrations should be present, which should prolong destruction of ozone by chlorine chemistry in the Arctic vortex during 1993. This is consistent with the MLS observations, as shown below.

The behaviour of LIMS ozone during 1978–79 is included as a baseline for a year when ozone loss in the lower stratosphere due to chlorine chemistry was negligible, because stratospheric chlorine abundances were lower by ~70% than at present. The delayed increase in vortex-averaged ozone in early winter 1978–79 and the brief ozone decrease at the end of January 1979 (Fig.

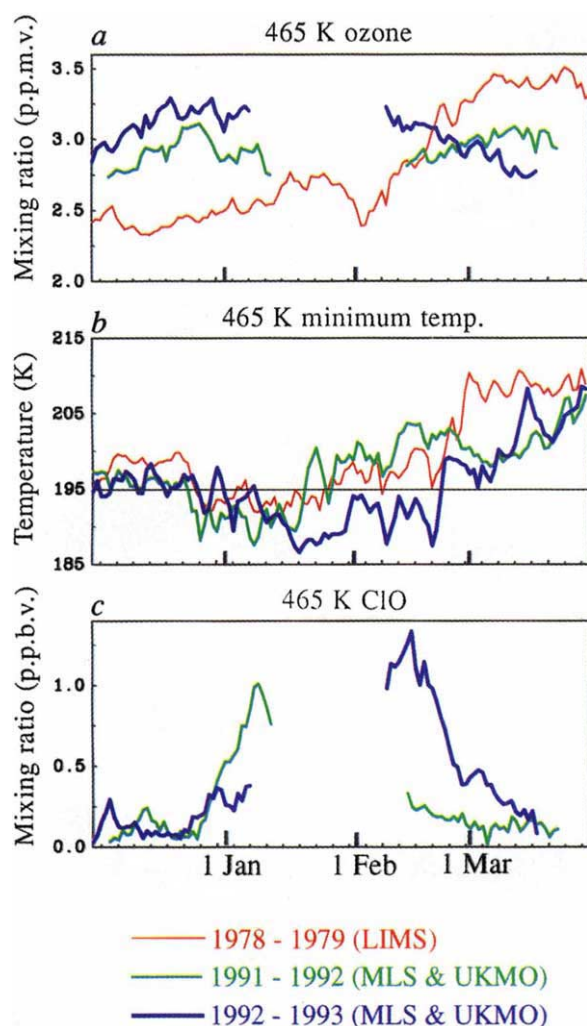
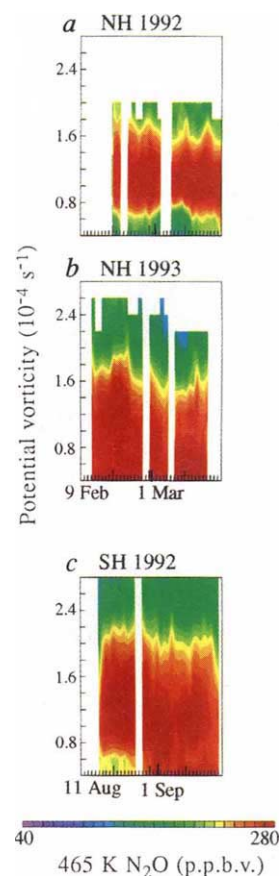


FIG. 3 *a* and *c*, Vortex-averaged mixing ratios, calculated by dividing the area integral of the mixing ratio inside a PV contour (the outermost contour in Figs 1 and 2), by the area enclosed by the contour, of ozone (*a*) and ClO (*c*) on the 465 K isentropic surface, for three Arctic winters; *b*, minimum temperatures (on the same isentropic surface) in the same region used for the vortex averages. Horizontal line in *b* indicates nominal threshold below which PSCs can form. Averaging MLS data as shown here improves the precision to ~0.05 p.p.m.v. for ozone and ~0.1 p.p.b.v. for ClO. The thin red line shows the 1978–79 winter (LIMS data), and the medium green and thick blue lines show the 1991–92 and 1992–93 winters, respectively (MLS ozone and ClO data, and UKMO temperatures). The gaps in MLS data are when it is observing the Southern Hemisphere. Using PV calculated from NMC rather than LIMS or UKMO does not significantly affect any results shown in Figs 3–6.

3*a*) suggest caution when interpreting recent decreases in vortex-averaged ozone as due to chemical rather than dynamical processes. The relatively low vortex ozone at 465 K in December 1978 does not necessarily imply a bias between LIMS and MLS observations, as vortex-averaged ozone was similar (within estimated measurement uncertainties^{17,20}) in early November in all three years. However, the dynamical evolution of the vortex was different in 1978–79. The winter observed by LIMS was relatively warm and the lower stratospheric vortex was much weaker, and therefore less isolated, than in either 1991–92 or 1992–93. The behaviour in 1978–79 was not atypical; in fact, a similar pattern, with maximum ozone near the vortex edge, was seen by MLS in the 1993–94 Arctic winter, when the lower stratospheric vortex was also relatively weak. As strong stratospheric warmings perturbed the late January and February 1979 vortex, strongest diabatic descent is expected then at the periphery of the vortex²⁷, leading to higher ozone there (for example, Fig. 1*a*). When highest ozone is near the outside edge of the vortex, it is not included in the average shown in Fig. 3*a*. Intrusions of low-latitude, ozone-poor air into the vortex²⁵ are also more common when it is weak, and rapid changes in LIMS ozone such as the brief decrease in late January 1979 can be due to horizontal transport¹⁸.

MLS observed enhanced ClO in the Arctic vortex from late December 1991 to mid-January 1992 (Fig. 3*c*), during the time of low temperatures¹⁰ (Fig. 3*b*). By mid-February, when MLS looked north again, ClO had decreased substantially and continued to decay through late February and early March. Vortex-averaged ozone decreased slightly at 465 K from early- to mid-January 1992 (Fig. 3*a*), consistent with vortex-averaged loss

FIG. 4 *a*–*c*, Time series of CLAES N₂O mixing ratios, averaged around a PV contour, as a function of PV, for 9 February (11 August) to 20 March (19 September) in the Northern Hemisphere, NH (Southern Hemisphere, SH). –PV is used in the Southern Hemisphere, where PV is negative. PV is scaled in 'vorticity units' (s⁻¹), where it is divided by a standard atmosphere value of the static stability^{13,43}. The bottom of the PV range shown ($0.4 \times 10^{-4} \text{ s}^{-1}$) is at low latitudes (~20°); PV values of $\sim(1.0\text{--}1.4) \times 10^{-4} \text{ s}^{-1}$ are representative of the vortex edge. The CLAES²¹ N₂O measurements have ~400 km horizontal resolution and ~2.5 km vertical resolution⁴⁴. Single profile precision and systematic error estimates in the lower stratosphere for this data version (no. 6) are ~20 p.p.b.v. r.m.s. and ~20%, respectively. N₂O mixing ratios >~210 p.p.b.v. are less reliable than lower values, due to lack of tangent point sensitivity⁴⁴. On the 465 K isentropic surface shown here, N₂O values are most reliable poleward of ~50° latitude. The averaging done here results in reliable trends for PV greater than $\sim(0.6\text{--}0.8) \times 10^{-4} \text{ s}^{-1}$. CLAES data are gridded by a simple interpolation of 24 h of data (providing full longitudinal coverage) to a regular latitude–longitude grid. Low N₂O mixing ratios at low PV in 1992 are artefacts caused by a Pinatubo-aerosol-related problem⁴⁴.



calculated using MLS-observed ClO^{10} . The ozone decrease in late January 1979, when ozone destruction by chlorine chemistry was not expected, is comparable to the early January 1992 decrease. But because the vortex was much stronger in 1992 than in 1979, ozone was more confined in 1992, and a comparable decrease caused by dynamics was thus not expected. Furthermore, on a timescale of a month, transport processes should cause an ozone increase during any active period, as seen in February 1979. The fact that vortex-averaged ozone did not increase following the strong warming in mid-January 1992 is circumstantial evidence for chemical ozone loss during that time, consistent with analyses of aircraft data²²⁻²⁴.

In the 1992-93 Arctic winter, MLS first detected enhanced ClO in early December, and vortex-averaged values at 465 K were ~ 0.5 p.p.b.v. by early January. There were no MLS Arctic observations from mid-January to mid-February 1993, but vortex minimum temperatures remained well below the PSC forma-

tion threshold, suggesting that ClO was enhanced throughout this period. When MLS looked north again in February, ClO mixing ratios were >1 p.p.b.v. in most of the vortex region. Temperatures remained below the PSC formation threshold until late February, after which ClO decreased.

Ozone at 465 K in the 1992-93 Arctic winter is expected to increase more rapidly during the warmings in late February and early March when diabatic descent is enhanced²⁷, similar to the

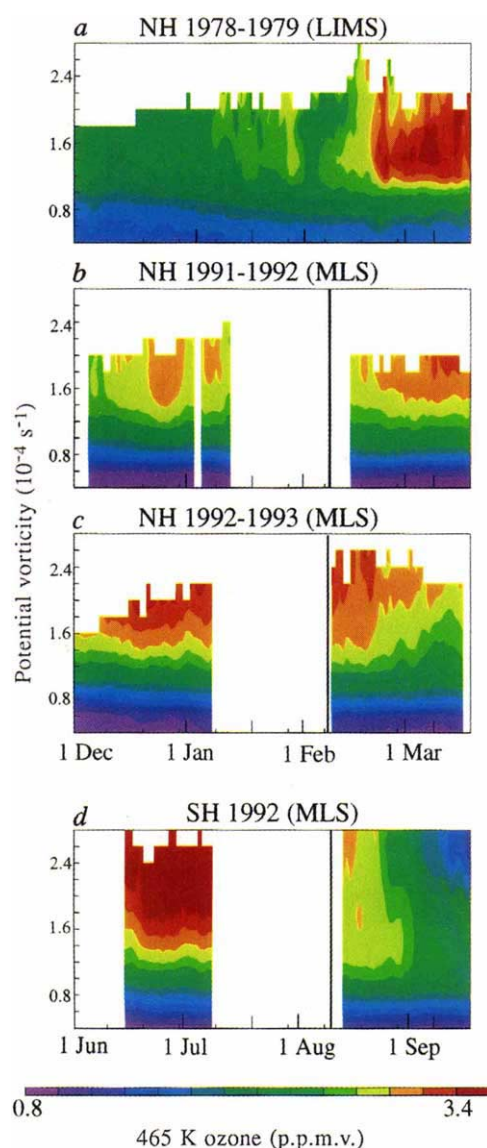


FIG. 5 a-d, Time series of ozone mixing ratios averaged around a PV contour, as a function of PV, for 1 December (1 June) to 20 March (19 September) in the Northern Hemisphere, NH (Southern Hemisphere, SH). Heavy vertical black lines indicate the time after which N_2O is shown in Fig. 4. Large gaps in the data are when MLS is looking at the hemisphere not shown.

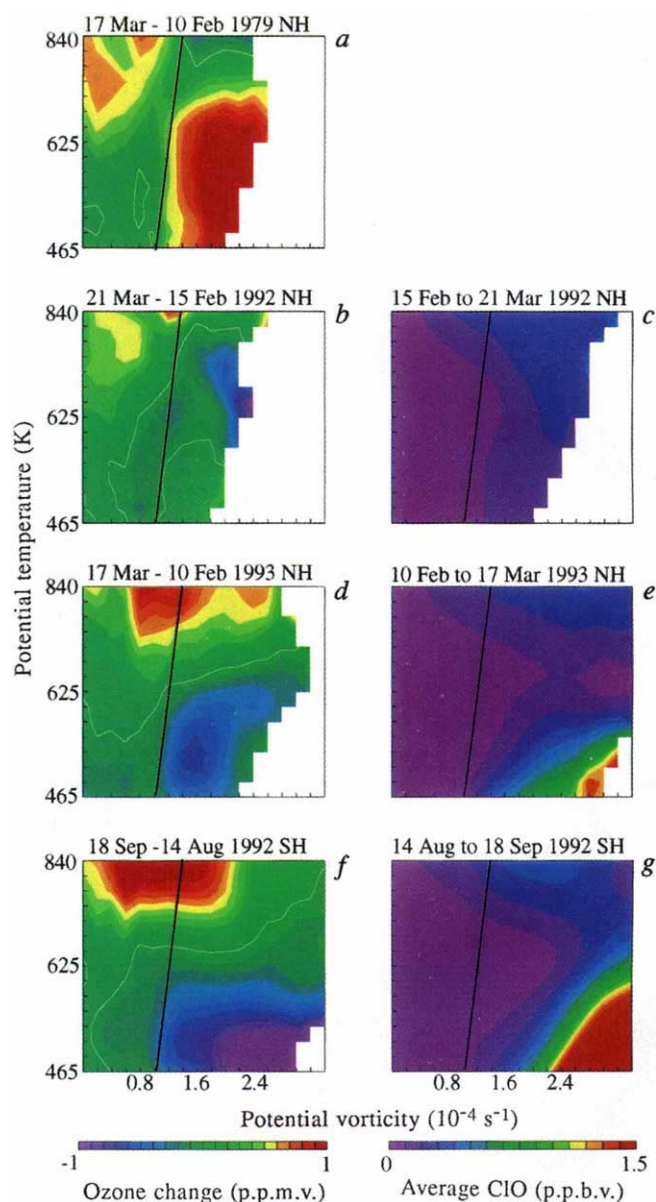


FIG. 6 Change in ozone mixing ratio (left-hand column), and time-averaged ClO (right-hand column), over 36 days (the approximate length of time for which MLS observed one hemisphere) in late winter, for three Arctic winters (no ClO data exist for 1979) and one Antarctic winter, as a function of PV and potential temperature (θ , ref. 31), in the lower to mid-stratosphere. White contour indicates no change. PV is scaled as in Figs 4 and 5, to give a similar range of values throughout the θ domain shown. Heavy black lines indicate PV typical of the vortex edge. Because the polar vortex weakens over the time period, especially in the Arctic, maximum PV values are less on the ending than on the beginning day. Ozone difference plots include all PV values present on both days differenced, whereas the ClO time-average plots include all PV values present on any day in the period. The ClO averaged at the highest PV values thus occurs earlier in the period.

behaviour seen in February 1979. Observed ozone increases gradually in the 1992–93 early winter, as expected. Vortex-averaged ozone is, however, the same in mid-February as in mid-January 1993, and decreases steadily between mid-February and mid-March. Between 10 February and 11 March, the average ozone mixing ratio decreases from approximately 3.2 p.p.m.v. to 2.6 p.p.m.v., $\sim 0.7\%$ per day. For comparison, vortex-averaged ozone in the Antarctic in August and September 1992 decreased at $\sim 1.4\%$ per day²⁸. Vortex-averaged ClO mixing ratios in the Antarctic in mid-August 1992 were comparable to mid-February 1993 Arctic values, but the duration of enhanced Arctic ClO was shorter^{10,28}, and Antarctic ClO remained enhanced later in the year when more sunlight was present. Calculations using MLS ClO like those done for January 1992¹⁰ give a vortex-averaged ozone loss over the February–March 1993 time period of $\sim 0.9\%$ per day due to chlorine (and bromine) chemistry. This calculated loss is consistent with the observations, particularly because ozone should increase in absence of chemical depletion.

Separation of chemical and dynamical effects

The behaviour of ozone is now examined in relation to long-lived tracers²⁹ to discriminate further between dynamical and chemical contributions to ozone change. As the Arctic vortex is so distorted and variable (Figs 1 and 2), zonal means obscure details of the flow, and average together values from inside and outside the vortex. A more useful average for diagnostic purposes is obtained by changing coordinates from latitude to PV (refs 30–33). Changes in the mixing ratio at constant PV and θ provide information on diabatic and chemical processes^{29–31,34}. Figure 4 shows time series of CLAES N_2O on the 465 K isentrope, averaged around PV contours for the Arctic late winters of 1992 and 1993, and the Antarctic late winter of 1992. Similar features are seen in plots of CLAES CH_4 , so only N_2O is shown here. The N_2O mixing ratio decreases toward the winter pole and decreases rapidly with height in the lower stratosphere, so diabatic descent would decrease N_2O along a PV contour on an isentropic surface³⁴. The N_2O shows a slight temporal decrease during February–March 1992 in the vortex at $PV \geq 1.4 \times 10^{-4} s^{-1}$ (PV scaled in vorticity units is used here—see Fig. 4 caption), indicating weak descent. A steeper decrease is seen in the middle of the February–March 1993 period, as expected from enhanced diabatic descent during stratospheric warmings²⁷. The N_2O level decreases slightly above the $1.4 \times 10^{-4} s^{-1}$ PV contour during August–September 1992 in the Antarctic, where minor warmings^{35,36} are confined to higher altitudes and do not significantly enhance diabatic descent in the lower stratosphere.

Figure 5 shows similar time series for ozone throughout the 1978–79, 1991–92 and 1992–93 Arctic winters, and the 1992 Antarctic winter. Ozone should increase where N_2O decreases if no chemical depletion occurs. Chemical processes were not expected to be important in the Arctic winter lower stratosphere in 1978–79; the general increase in ozone is consistent with changes expected from diabatic descent. In 1991–92, ozone at 465 K increases slightly in December and in late February, consistent with the transport processes indicated by concurrent changes in N_2O . In stark contrast, the evolution of ozone in the 1992 Antarctic late winter is very different from that expected from dynamical variations as diagnosed from N_2O ; ozone decreases rapidly, beginning near the centre of the vortex.

Although less dramatic, the evolution of 465 K ozone during the 1992–93 Arctic winter resembles that in the Antarctic 1992 winter. Arctic ozone slowly increases in December 1992, as does Antarctic ozone in June 1992. But in mid- to late-February 1993, Arctic ozone decreases rapidly well inside the vortex. Comparison with dynamical changes expected from examination of N_2O (Fig. 4b) indicates that the observed ozone evolution in February–March 1993 is inconsistent with the effects of transport alone. This provides direct observational evidence for chemical depletion of ozone in the Arctic during February–March 1993.

Figure 6 compares the vertical extent of enhanced ClO with ozone changes in late winter. The 1979 ozone (Fig. 6a) shows behaviour expected for no chemical loss during a dynamically active period: a large increase inside the polar vortex throughout the lower stratosphere. Figure 6f and g, for 1992 in the Antarctic, shows the behaviour expected when chemical loss dominates: ozone decreased below ~ 650 K (~ 27 km) throughout the vortex, in the region of enhanced ClO.

The level of ClO in the Arctic lower stratosphere during February–March 1992 was low compared with that observed when temperatures are below the PSC formation threshold. Ozone in February–March 1992 below ~ 650 K increased slightly inside the vortex, consistent with weak diabatic descent at this time. An ozone decrease of $\sim 10\%$ occurred in 1992 between 600 and 700 K at highest PV values; a decrease of $\sim 7\%$ was seen in this region in 1993. First-order estimates of gas-phase ozone destruction reaction rates from catalytic cycles involving NO_2 (from CLAES) and ClO suggest that much of the ozone difference between 1992 and 1993 near 650 K could have been caused by the somewhat larger abundances of these radicals in 1992.

Enhanced Arctic ClO in 1993 extended to near the vortex edge and vertically to ~ 580 K. Although Arctic ClO abundances in mid-February 1993 were comparable to Antarctic values in mid-August 1992, the time average was less in the Arctic because ClO decreased in early March (Fig. 3c), and thus the period of enhancement was less. Ozone decreased in 1993 (Fig. 6d) throughout the vortex below ~ 600 K, in the general region of enhanced ClO.

Implications for Arctic ozone loss

We have demonstrated above that the $\sim 20\%$ decrease in Arctic lower-stratospheric ozone in February and March 1993 is inconsistent with changes expected from transport alone during this period. The decrease seems to be consistent with the observed ClO and associated depletion by chlorine chemistry.

Although ozone in the lower stratospheric vortex decreased substantially in late winter 1993, MLS observations showed no significant decrease in column ozone for the same time²⁰. Ozone increases rapidly at this time in the Arctic polar mid-stratosphere, compared with much smaller increases in the Antarctic mid-stratosphere³⁵. Higher altitudes thus contribute more significantly to the column in the Arctic, masking the lower-stratospheric ozone decrease there. Comparison with a 30-year record of Dobson data³⁷ shows that the change in Arctic column ozone from MLS between February and March 1993 is less of an increase than commonly observed, but not the smallest increase on record.

Arctic lower-stratospheric temperatures were low enough for PSC formation for ~ 1 month during 1991–92^{10,13}, ending in late January. They were low enough for nearly 2 months during 1992–93, until late February, when more sunlight reaches the polar regions and sustains the ozone destruction cycle over larger areas. This year (1993–94), Arctic lower-stratospheric temperatures hovered near and slightly below the PSC formation threshold from December 1993 to late February 1994, when they dropped well below it for ~ 2 weeks; MLS measured substantial enhancement of ClO in the Arctic vortex in late February/early March 1994. The Arctic lower stratosphere in the past 16 years has been sufficiently cold for PSC formation for as long as $2\frac{1}{2}$ months, and as late as mid-March¹⁵. It is therefore expected that there will be years in the near future, while stratospheric chlorine levels continue to increase, when ozone depletion in the Arctic lower stratosphere will equal or exceed that in 1993. □

Received 21 February; accepted 11 July 1994.

- Farman, J. C., Gardiner, B. G. & Shanklin, J. D. *Nature* **315**, 207–210 (1985).
- Solomon, S. *Nature* **347**, 347–354 (1990).
- Anderson, J. G., Brune, W. H. & Proffitt, M. H. *J. geophys. Res.* **94**, 11465–11479 (1989).
- de Zafra, R. L. et al. *Nature* **328**, 408–411 (1987).
- Solomon, S., Garcia, R. R., Rowland, F. S. & Wuebbles, D. J. *Nature* **321**, 755–758 (1986).
- Molina, M. J., Tso, T. L., Molina, L. T. & Wang, F. C. Y. *Science* **238**, 1253–1257 (1987).

7. Brune, W. H., Toohey, D. W., Anderson, J. G. & Chan, K. R. *Geophys. Res. Lett.* **17**, 505–508 (1990).
8. Webster, C. R. et al. *Science* **261**, 1130–1134 (1993).
9. Toohey, D. W. et al. *Science* **261**, 1134–1136 (1993).
10. Waters, J. W. et al. *Nature* **362**, 597–602 (1993).
11. Andrews, D. G. *Pure appl. Geophys.* **130**, 213–232 (1989).
12. Schoeberl, M. R., Lait, L. R., Newman, P. A. & Rosenfield, J. E. *J. geophys. Res.* **97**, 7859–7882 (1992).
13. Manney, G. L. & Zurek, R. W. *Geophys. Res. Lett.* **20**, 1275–1278 (1993).
14. Hartmann, D. L. et al. *J. geophys. Res.* **94**, 16779–16796 (1989).
15. Northern Hemisphere Winter Summary, 1992–1993, *Selected Indicators of Stratospheric Climate* (NOAA, Climate Analysis Centre, National Meteorological Center, Washington DC, 1993).
16. Hoskins, B. J., McIntyre, M. E. & Robertson, A. W. *Q. Jl R. met. Soc.* **111**, 877–946 (1985).
17. Remsberg, E. E. et al. *J. geophys. Res.* **89**, 5161–5178 (1984).
18. Leovy, C. B. et al. *J. atmos. Sci.* **42**, 230–244 (1985).
19. Barath, F. T. et al. *J. geophys. Res.* **98**, 10751–10762 (1993).
20. Froidevaux, L. et al. *J. atmos. Sci.* (in the press).
21. Roche, A. E. et al. *J. geophys. Res.* **98**, 10763–10775 (1993).
22. Browell, E. V. et al. *Science* **261**, 1155–1158 (1993).
23. Proffitt, M. H. et al. *Science* **261**, 1150–1154 (1993).
24. Salawitch, R. J. et al. *Science* **261**, 1146–1149 (1993).
25. Plumb, R. A. et al. *J. geophys. Res.* **99**, 1089–1105 (1994).
26. Newman, P. A. et al. *Science* **261**, 1143–1146 (1993).
27. Manney, G. L. et al. *Geophys. Res. Lett.* **21**, 813–816 (1994).
28. Waters, J. W., Froidevaux, L., Manney, G. L., Read, W. G. & Elson, L. S. *Geophys. Res. Lett.* **20**, 1219–1222 (1993).
29. Schoeberl, M. R. et al. *Geophys. Res. Lett.* **17**, 469–472 (1990).
30. Butchart, N. & Remsberg, E. E. *J. atmos. Sci.* **43**, 1319–1339 (1986).
31. Schoeberl, M. R. et al. *J. geophys. Res.* **94**, 16815–16845 (1989).
32. Lait, L. R. et al. *Geophys. Res. Lett.* **17**, 521–524 (1990).
33. Schoeberl, M. R. & Lait, L. R. in *EOS NATO Summer School* (eds Visconti, G. & Gilie, J.) 419–431 (North-Holland, Amsterdam, 1992).
34. Haynes, P. H. & McIntyre, M. E. *J. atmos. Sci.* **44**, 828–841 (1987).
35. Manney, G. L. et al. *Geophys. Res. Lett.* **20**, 1279–1282 (1993).
36. Fishbein, E. F. et al. *Geophys. Res. Lett.* **20**, 1255–1258 (1993).
37. *Rep. int. Ozone Trends Panel 1988* (Global Ozone Res. Monitoring Proj. Rep. No. 18, WMO, Geneva, 1988).
38. Elson, L. S. & Froidevaux, L. *J. geophys. Res.* **98**, 23039–23049 (1993).
39. Swinbank, R. & O'Neill, A. *Mon. Weath. Rev.* **122**, 686–702 (1994).
40. Randel, W. J. *J. atmos. Sci.* **44**, 3097–3120 (1987).
41. Finger, F. G. et al. *Bull. Am. met. Soc.* **74**, 789–799 (1993).
42. Manney, G. L., Zurek, R. W., Gelman, M. E., Miller, A. J. & Nagatani, R. *Geophys. Res. Lett.* (submitted).
43. Dunkerton, T. J. & Delisi, D. P. *J. geophys. Res.* **91**, 1199–1208 (1986).
44. Kumer, J. B., Mergenthaler, J. L. & Roche, A. E. *Geophys. Res. Lett.* **20**, 1239–1242 (1993).

ACKNOWLEDGEMENTS. We thank our MLS and CLAES colleagues for their contributions to the success of those experiments, T. Luu for data management, P. Newman for supplying routines that were adapted to calculate PV, R. A. Plumb, M. R. Schoeberl, W. J. Randel, C. B. Leovy and J. R. Holton for helpful comments or discussions. The work at the Jet Propulsion Laboratory, California Institute of Technology was sponsored by NASA and the work at Edinburgh University by SERC.

A protein complex required for signal-sequence-specific sorting and translocation

Brigitte Wiedmann*, Hideaki Sakai, Terri A. Davis & Martin Wiedmann

Cellular Biochemistry and Biophysics Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA

We have purified a nascent-polypeptide-associated complex (NAC) which prevents short ribosome-associated nascent polypeptides from inappropriate interactions with proteins in the cytosol. NAC binds nascent-polypeptide domains emerging from ribosomes unless a signal peptide is fully exposed. Depletion of cytosolic proteins (including NAC) from ribosomes carrying nascent polypeptides allows the signal recognition particle (SRP) to crosslink to polypeptides irrespective of whether or not they contain signal peptides. In the absence of cytosol, proteins lacking signal peptides can be mistranslocated into the endoplasmic reticulum *in vitro*, albeit with low efficiency. Readdition of NAC restores the specificity of SRP and fidelity of translocation.

SEVERAL proteins are known to interact with ribosome-associated nascent chains (hereafter referred to simply as nascent chains). Proteins such as the chaperones or modifying enzymes (for example, signal peptidase, oligosaccharyl transferase or protein disulphide isomerase) can interact with both nascent chains and polypeptides that have been released from the ribosome^{1–5}. Here we describe the identification and purification of a nascent-polypeptide-associated complex (NAC), a heterodimeric protein which binds to short nascent chains but not to released chains, with the exception of those with fully exposed signal peptides. Using *in vitro* assay systems lacking most cytosolic factors (including NAC) that bind nascent chains in a salt-sensitive manner, we show that signal recognition particle (SRP) can interact with proteins lacking signal peptides. SRP, which acts as a shuttle between the cytosol and the endoplasmic reticulum (ER) membrane, is the only other known 'non-ribosomal' factor

which interacts exclusively with nascent, as opposed to released, chains. When bound to signal peptides, SRP, by interacting with its receptor at the ER membrane, mediates the targeting and transport of nascent chains to and across the ER membrane^{2,5}. We now show that SRP can mediate the targeting to and translocation across the ER membrane of proteins lacking signal peptides.

Readdition of purified NAC to cytosol-depleted ribosomes both restores the specificity of SRP binding and prevents mistargeting and therefore inappropriate translocation. We hypothesize that, as NAC is likely to be one of the first non-ribosomal proteins with which a nascent chain makes contact, it functions in part to shield non-signal peptide regions of nascent chains from potentially promiscuous interactions with SRP.

Identification of NAC

To identify proteins interacting with nascent chains, we translated truncated messenger RNAs encoding different proteins and lacking stop codons *in vitro* to produce stable ribosome-associated nascent-chain intermediates of defined length(s)⁶.

* Present address: Mount Sinai Medical Center, Department of Cell Biology/Anatomy, New York 10029, USA.

Additionally, instead of lysine, lysyl residues derivatized with the photoactivatable crosslinker TDBA (4-(3-trifluoromethyl-diazirino)benzoic acid) were incorporated at known positions. These nascent chains bearing photoactivatable probes are substrates for nascent-chain-binding proteins, and photoactivation causes covalent attachment of interacting partners. The highly reactive carbenes produced on photoactivation of this very short probe have a half-life of about 1 ns and can react with any chemical bond⁷, thus allowing control of spatial and temporal conditions of crosslinking. Following sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and fluorography, the size of the crosslinked protein is estimated by subtracting the relative molecular mass (M_r) of the nascent chain from the size of the photoadduct⁸. The nascent chains used here are: peroxisomal firefly luciferase (fLuc), mitochondrial pre-F¹ ATPase β -subunit (pF¹ β), chloramphenicol acetyl transferase (CAT), mitochondrial pre-cytochrome oxidase IV subunit (pCOXIV), bovine preprolactin (pPL) and a described signal peptide mutant of pPL (pPL-M) that is not translocated⁹. The number of translated amino acids (aa) precedes the name of each nascent chain. Figure 1a shows that nascent chains of 52aa fLuc, 77aa fLuc, 55aa pPL-M and 51aa pF¹ β can be crosslinked to several proteins present in the wheat-germ cytosol, with a crosslink to a protein of $M_r \sim 33,000$ (33K) being predominant (see Fig. 1a legend for positions of lysines and methionines).

To determine whether or not the crosslinked protein is a core ribosomal protein, we treated the nascent chains with increasing salt concentrations before sedimentation through salt/sucrose cushions and subsequent irradiation. As shown in Fig. 1b (lanes 1–5), the 33K protein is bound to ribosome-associated nascent chains in a salt-sensitive manner. Puromycin release of the nascent chain before irradiation abolished the crosslink, suggesting that the 33K protein only binds polypeptides in the context of the ribosome. Puromycin treatment after irradiation released the crosslinked product from the ribosome (not shown, but see Fig. 2b), supporting the view that the 33K protein is not a core ribosomal protein.

To investigate whether this protein binds reversibly, nascent chains were extracted with high salt and isolated by sedimentation through a salt/sucrose cushion. The 33K protein was crosslinked when cytosol from wheat germ (Fig. 1b, lanes 6, 7), bovine brain (lanes 8, 9) and dog pancreas (lanes 10, 11) was added to the bare nascent chains before irradiation. Therefore, the 33K protein is salt-extractable, can rebinding after extraction and is present in a variety of species and tissues.

Purification of NAC

We used salt-stripped 52aa fLuc nascent chains as the substrate for purifying the 33K protein from bovine brain cytosol (Fig. 2a). Throughout our purification, a 33K protein co-purified with a 21K protein which is also evident but not dominant in crude cytosol (see Fig. 1a). These proteins, named α NAC and β NAC, form a heterodimeric complex (H.S. *et al.*, manuscript in preparation). Both subunits could be crosslinked to the nascent chain (Fig. 2b and c). Puromycin addition before crosslinking prevented purified NAC from binding to nascent chains (Fig. 2b) and its addition after crosslinking released crosslinked NAC along with the nascent chain (data not shown).

Peptide sequence data from purified α - and β NAC and subsequent complementary DNA cloning of the human genes indicated the following homologies: α NAC is homologous to three overlapping DNA fragments of unknown function which encode all but the amino-terminal 23 amino acids of α NAC (GenBank accession numbers D12194, Z28479 and Z25168)¹⁰. β NAC is identical to BTF3b, a protein of unknown function¹¹.

To prove that the proteins we purified are indeed the ones that give rise to the crosslinks, peptide-specific antibodies to the deduced carboxy-terminal amino-acid sequences of α - and β NAC were generated and used to immunoprecipitate the crosslinked nascent chains (Fig. 2c).

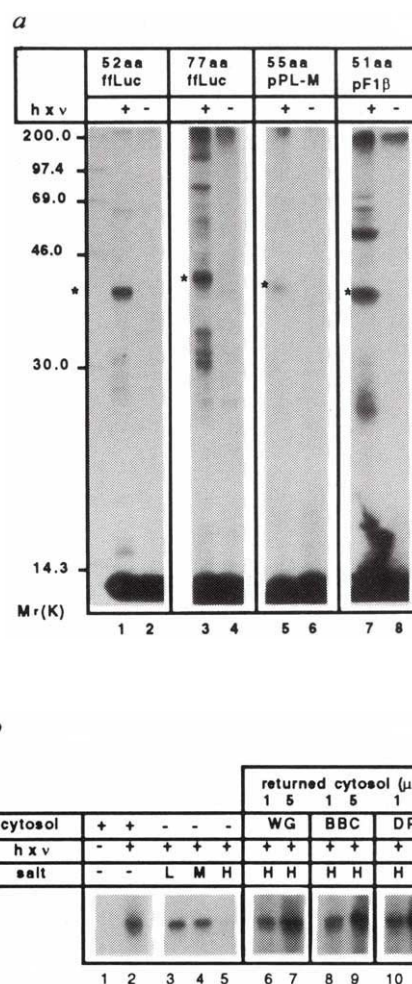


FIG. 1 A 33K protein present in the cytosol of various tissues can bind to short nascent chains in a salt-sensitive manner. **a**, Crosslinking in wheat-germ lysate. 'h x v' indicates irradiation and the asterisk marks the crosslinking product between the 33K protein and the nascent chains. **b**, The 33K protein binds in a salt-sensitive manner and is present in various cell types. The crosslink between 52aa fLuc nascent chains and the 33K protein (lane 2) is abolished when nascent chains were adjusted to 500 mM potassium acetate (H) and sedimented through sucrose cushions. Return of cytosol from wheat germ (lanes 6, 7), bovine brain (lanes 8, 9) or dog pancreas (lanes 10, 11) leads to the reappearance of the crosslink.

METHODS. **a**, *In vitro* translation of truncated mRNAs and photo-crosslinking are as described elsewhere^{6,8}. Samples were treated with RNase A, precipitated with TCA and visualized by fluorography after SDS-PAGE. The positions of lysines and methionines in the proteins are given:

	Lysines	Methionines
77aa fLuc	5, 8, 9, 28, 31, 67	1, 30, 58, 66
86aa pPL, 86aa pPL-M	4, 9, 72, 78	1, 54, 66, 83
96aa pCOX IV	12, 28, 32, 52, 75, 85, 92	1, 95
51aa pF ¹ β	16, 19, 30, 45	1, 33
76aa CAT	3, 4, 19, 46, 49, 50, 54	1, 66, 75

b, After translation, aliquots were adjusted to 150 mM (L), 250 mM (M) or 500 mM (H) potassium acetate and spun through 0.5 M sucrose cushions in TBB (translation mixture lacking lysate, radiolabel and mRNA) of the same potassium acetate concentration (20 min, TL120.1 rotor, Beckman; 120,000 r.p.m., 1 vol. sample over 3 vols cushion). Pellets were resuspended in a half-volume of TBB. Aliquots (7 μ l) were either irradiated immediately and processed as described above (lanes 3–5) or processed after incubation for 5 min on ice with 1 or 5 μ l cytosol (WG: gel filtrated, 20,000g supernatant, 30 mg ml⁻¹; BBC: gel filtrated, 100,000g supernatant, 25 mg ml⁻¹; DPC: 100,000g supernatant, 15 mg ml⁻¹).

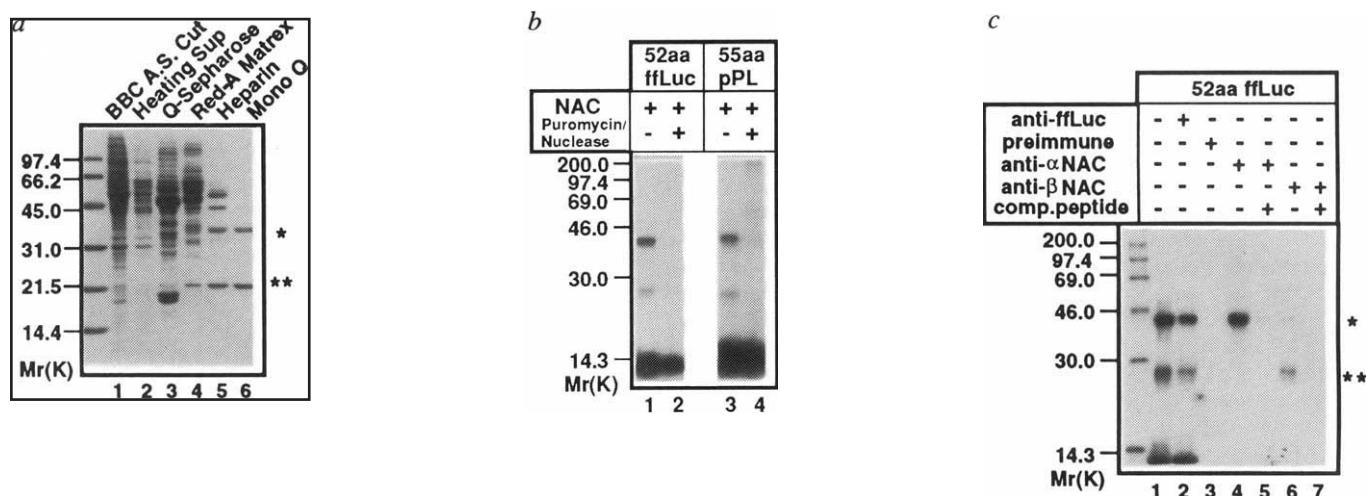


FIG. 2 Purification and partial characterization of NAC. **a**, Protein pattern at different stages of purification after SDS-PAGE and Coomassie blue staining. Lane 1, bovine brain cytosol, dialysed with 30–60% ammonium sulphate fraction, 16 μ g; lane 2, supernatant after heating, 6 μ g; lane 3, Q-Sepharose FF active fraction, 15 μ g; lane 4, Red-A Matrex active fraction, 12 μ g; lane 5, Heparin-HiTrap active fraction, 4 μ g; lane 6, Mono Q active fraction, 2 μ g. Two proteins, α NAC and β NAC, were copurified by this procedure. Crosslinks to NAC are indicated in all figures by one (α NAC) or two (β NAC) asterisks. **b**, Purified NAC binds only to nascent-chain ribosome complexes. When nascent chains were released from the ribosomes, no crosslink to NAC is observed (lanes 2 and 4). **c**, The purified α NAC and β NAC proteins are identical to the crosslinked proteins. 52aaffLuc nascent chains were crosslinked to purified NAC and then either subjected to SDS-PAGE (lane 1) or immunoprecipitated with firefly luciferase (lane 2), preimmune (lane 3), α NAC (lanes 4 and 5) or β NAC (lanes 6 and 7) sera. The specificity of each antibody was verified by competition with antigenic peptide (lanes 5 and 7).

METHODS. **a**, Bovine brain was homogenized in 2 vols 250 mM sucrose in buffer A (25 mM Tris-HCl pH 8.0, 50 mM KCl, 1 mM dithiothreitol (DTT), 1 mM EGTA, 5 mM $MgCl_2$, 1 mM phenylmethyl sulphonyl fluoride (PMSF)). After centrifugation (Ti45, Beckman; 40,000 r.p.m., 1.5 h) a dialysed 30–60% ammonium sulphate fraction of the supernatant was heated for 5 min at 70 °C and then spun as before, but for 2 h. The cleared supernatant was applied to a Q-Sepharose FF column (Pharmacia). The activity in each fraction was measured by crosslinking to

salt-stripped 52aaffLuc nascent chains, which gave major photoadducts of 42K and 27K. Fractions over 240–430 mM KCl were combined and then applied to a Red-A Matrex column (Amicon). After washing with 300 mM KCl in buffer B (20 mM HEPES pH 7.5, 1 mM DTT, 1 mM PMSF, 25% glycerol), the activity was eluted with 1 M KCl in buffer B. The eluent was diluted 10-fold with buffer B and applied to a prepacked Heparin-HiTrap column (Pharmacia). The eluent at 400 mM KCl was diluted 5-fold with buffer B, applied to a Mono-Q column (Pharmacia) and eluted with a linear gradient of 100–500 mM KCl in buffer B. A single protein peak was obtained at 320 mM KCl concentration which shows a single protein band in native PAGE (not shown) and two bands of 33K and 21K in SDS-PAGE, as shown in lane 6. From 800 g of bovine brain, 750 μ g of purified NAC was obtained. **b**, Salt-stripped nascent chains of 52aaffLuc or 55aapPL were treated for 10 min at 30 °C with 1 mM puromycin and 100 μ g ml^{-1} RNase A (lanes 2 and 4) or mock-treated (lanes 1 and 3). Purified NAC (100 nM) was allowed to bind for 5 min at 26 °C before irradiation. **c**, Immunoprecipitation. 52aaffLuc–NAC crosslink adducts were treated with RNase A (100 μ g ml^{-1}) for 10 min at 30 °C to digest peptidyl-transfer RNAs. The samples were adjusted to 1% SDS and immunoprecipitated by a standard procedure¹⁷. α NAC and β NAC antibodies were raised in rabbits against synthetic peptides corresponding to the C-terminal sequence of a human expressed sequence tag (HEST; GenBank accession number D12194), RALKNSNDIVNAIMELTM-COOH, and of BTF3b (GenBank accession number X53281), ENFDEASKNEAN-COOH, respectively.

NAC prevents inappropriate SRP crosslinking

Because the first lysine in the 51aapF β nascent chain is 16 amino acids from the N-terminus and this lysine crosslinks to the 33K protein (Fig. 1a) and to purified NAC (not shown), we conclude that NAC can bind regions of nascent chains as near as 35 amino acids from the peptidyl transferase. Thus, NAC is very near to where the nascent chain leaves the ribosome (exit site). The SRP 54K protein has previously been shown to bind to the signal sequence when near the exit site^{12–14}.

We have shown that in unfractionated cytosol, nascent chains lacking signal peptides primarily crosslink to NAC whereas proteins with signal peptides predominantly crosslink to SRP (see Figs 1a and 3a). The identity of the crosslink to SRP 54K was established by immunoprecipitation (not shown). In the case of a very short preprolactin nascent chain (55aapPL), we observed crosslinks to both SRP and NAC (Fig. 3a, lane 13). It has been shown^{13,14} that SRP-arrested preprolactin nascent chains of 70 amino acids crosslink solely to SRP 54K. To determine directly whether NAC and SRP have different specificities, we prepared cytosol-depleted nascent chains of 86aapPL, 77aaffLuc or 55aapPL and added SRP to each. Surprisingly, SRP crosslinked to all nascent chains regardless of whether or not they harboured a functional signal peptide (Fig. 3a, lanes 4,

9, 14; see also Fig. 3b). If cytosol was not depleted, SRP crosslinked only to signal peptide containing nascent chains (Fig. 3a, lanes 3, 13). To exclude the possibility that our observation resulted from using a heterologous system (wheat-germ ribosomes and canine SRP), we repeated these experiments in the rabbit reticulocyte lysate translation system (Fig. 3b). Again, SRP can be crosslinked to all the cytosol-depleted nascent chains.

When bound to ribosomes alone, SRP can be extracted with high salt¹⁵. In contrast, when bound to a signal peptide in the context of the ribosome, SRP is resistant to salt extraction¹⁶. Because crosslinking SRP to non-signal peptides may measure proximity rather than binding, we reasoned that if SRP is truly bound to a non-signal peptide, its binding should be resistant to high salt. The experiment in Fig. 3c shows this to be the case. Nascent chains of 86aapPL and 77aaffLuc were assembled and cytosol was removed (Fig. 3c, '1x'). After incubation with SRP, an aliquot of each reaction was irradiated. The remainder of the sample was extracted with high salt a second time and the bound SRP recovered by sedimentation of nascent-chain complexes. Because the amount of SRP that could be crosslinked to the nascent chain was not diminished, we conclude that SRP does interact with non-signal peptides and is not merely crosslinked

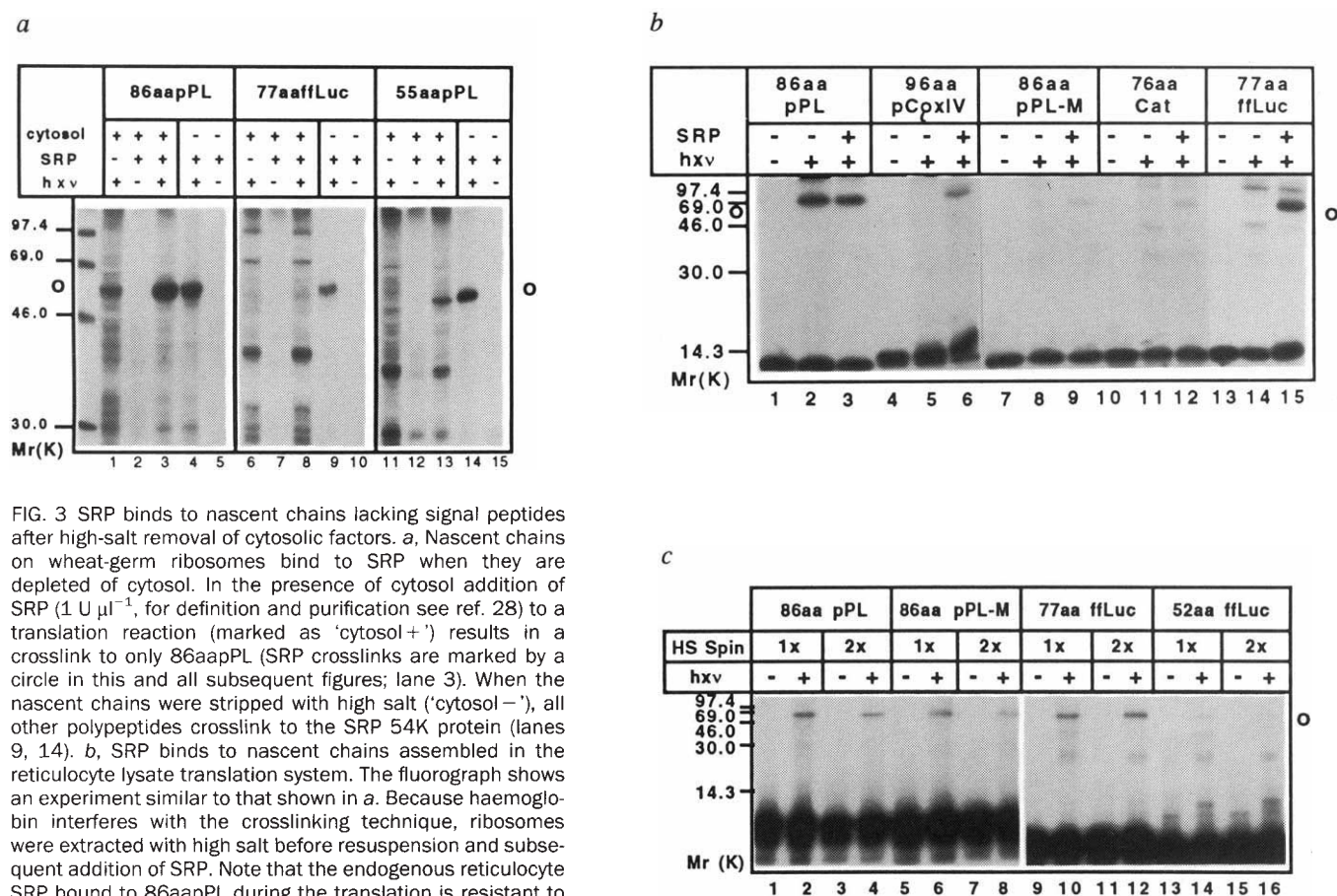


FIG. 3 SRP binds to nascent chains lacking signal peptides after high-salt removal of cytosolic factors. **a**, Nascent chains on wheat-germ ribosomes bind to SRP when they are depleted of cytosol. In the presence of cytosol addition of SRP ($1 \text{ U } \mu\text{l}^{-1}$, for definition and purification see ref. 28) to a translation reaction (marked as 'cytosol+') results in a crosslink to only 86aapPL (SRP crosslinks are marked by a circle in this and all subsequent figures; lane 3). When the nascent chains were stripped with high salt ('cytosol-'), all other polypeptides crosslink to the SRP 54K protein (lanes 9, 14). **b**, SRP binds to nascent chains assembled in the reticulocyte lysate translation system. The fluorograph shows an experiment similar to that shown in **a**. Because haemoglobin interferes with the crosslinking technique, ribosomes were extracted with high salt before resuspension and subsequent addition of SRP. Note that the endogenous reticulocyte SRP bound to 86aapPL during the translation is resistant to salt extraction (lane 2). **c**, SRP binding to ribosomes carrying both signal- and non-signal-peptide-containing nascent chains is salt resistant. This fluorograph shows that SRP bound to salt-stripped nascent chains ('1x', lanes 2, 6, 10, 14) is not removed by a second high-salt treatment ('2x', lanes 4, 8, 12, 16). Because varying amounts of

nascent chains were released from the ribosome during the second extraction, as assessed by scintillation counting, we loaded a correspondingly decreased amount of the '1x' samples so that the amount of nascent chains is identical in '1x' and '2x' samples.

to nascent chains by virtue of being in the right place at the wrong time.

As shown in Fig. 4a restoration of cytosol to stripped nascent chains of 52aaffLuc re-establishes the specificity of SRP crosslinking. Figure 4b shows that in the absence of added NAC, SRP 54K protein crosslinked to 52aaffLuc (non-signal peptide; lane 1). However, this crosslink diminished in a dose-dependent manner when increasing amounts of NAC were added and the NAC crosslink increased (Fig. 4b, lanes 2–5). The same phenomenon was observed with 86aapPL-M (carrying a mutant, inactive signal peptide; Fig. 4b, lanes 6–10). We have determined that NAC is present in different cytosols in the micromolar range (not shown) and therefore that the concentrations used represent cellular concentrations of NAC. As the SRP concentration was 20 nM ($\sim 1 \text{ U } \mu\text{l}^{-1}$), an approximately 10-fold molar excess of NAC is needed for half-maximal-competition (Fig. 4b, c). Thus, for nascent chains lacking functional signal peptides, SRP binding can be prevented by adding NAC. In contrast, when the experiment used 86aapPL, the SRP 54K protein crosslink was not diminished (Fig. 4b, lanes 11–15) and a crosslink to NAC was barely detectable. Figure 4c is a quantification of the same type of experiment shown in Fig. 4b but includes analysis of 55aapPL and 55aapPL-M. As would be predicted from the results shown in Fig. 3a, NAC also prevented SRP from binding to 55aapPL. With the exception of 86aapPL, half-maximal inhibition of SRP crosslinking occurred at NAC concentrations of 200–500 nM, all below the cellular concentrations of NAC.

Because the inappropriate binding of SRP to ribosome-associated non-signal peptides was resistant to high salt extraction (Fig. 3c), we imagined that NAC's ability to restore SRP specificity is not merely a result of adding bulk protein to the isolated ribosome/nascent-chain complexes. In this case, the inappropriate SRP crosslinking should persist on addition of a NAC-depleted cytosol or bovine serum albumin (BSA). Stripped 52aaffLuc nascent-chain complexes were assembled in the reticulocyte translation system and incubated with SRP as indicated in Fig. 4d. Buffer, complete bovine brain cytosol, NAC-depleted bovine brain cytosol or BSA were returned before crosslinking. In all cases where NAC was depleted, the SRP crosslink was evident (Fig. 4d, lanes 2, 4, 7). When purified NAC (Fig. 4d, lane 6), complete cytosol (Fig. 4d, lane 3) or NAC and depleted cytosol (Fig. 4d, lane 5) were returned, the SRP crosslink was abolished and the NAC crosslink increased. NAC was depleted by heparin chromatography, and only 10% of cytosolic protein, but >95% of NAC, bound.

NAC prevents inappropriate ER translocation

To test whether SRP bound to nascent chains lacking signal peptides could mediate targeting and functionally engage the SRP receptor, salt-stripped nascent-chain complexes were assembled and incubated with SRP. After binding of SRP, SRP-depleted ER membranes (ER membranes washed with EDTA and high salt) were added and samples irradiated. In the absence of membranes, the 86aapPL–SRP 54K protein crosslink was

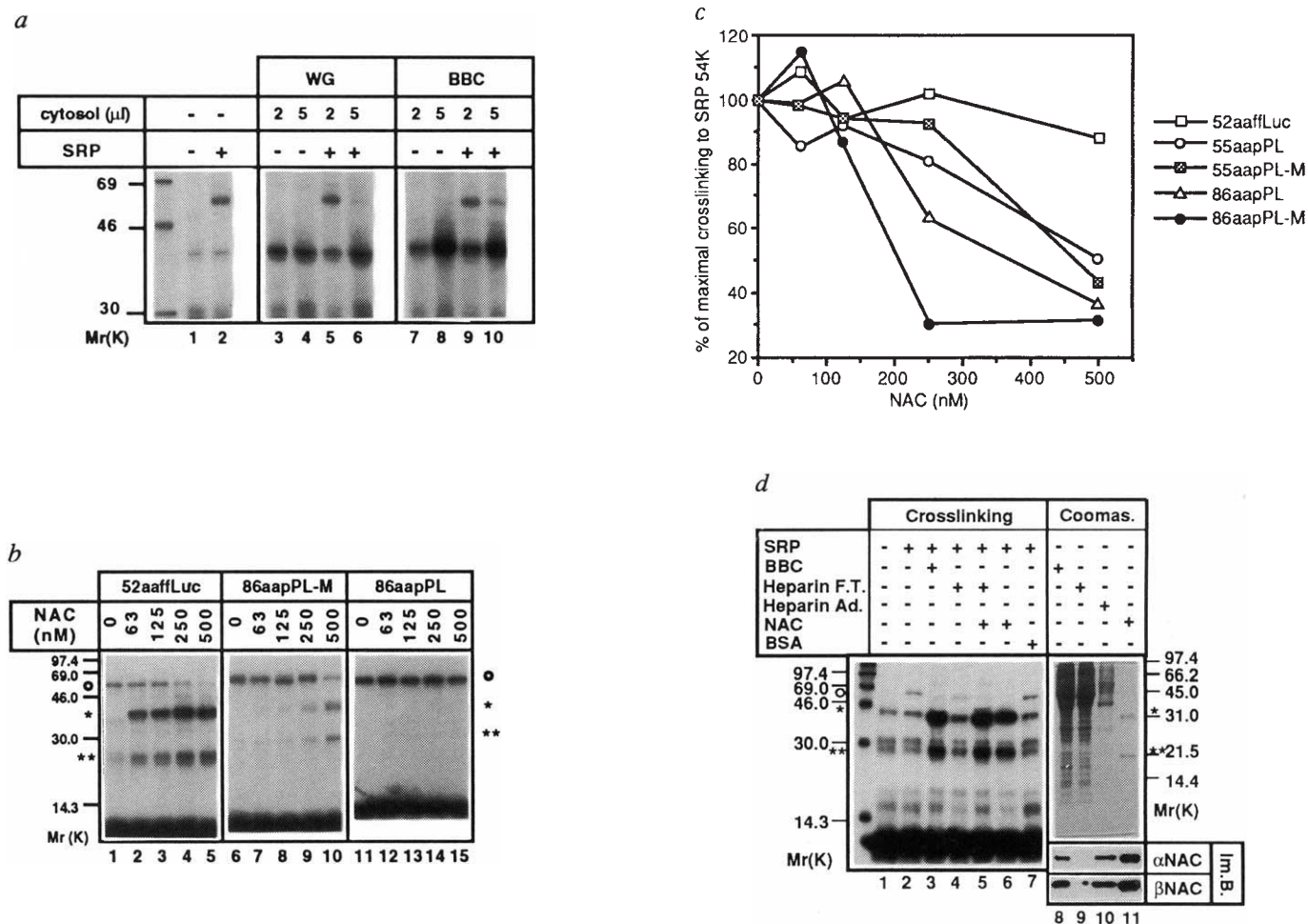


FIG. 4 Cytosolic proteins prevent the binding of SRP to nascent chains lacking signal sequences. **a**, SRP does not bind to 52aaffLuc when cytosol is returned to isolated nascent chains. Salt-stripped 52aaffLuc nascent chains were divided into ten 7-μl aliquots. SRP was allowed to bind for 5 min on ice (samples in lanes 2, 5, 6, 9, 10) and then cytosol (for concentration see Fig. 1b) was added (lanes 3–10). All samples were irradiated and processed as described for Fig. 1a. **b**, Purified NAC prevents SRP from binding to non-signal sequences. Salt-stripped nascent chains of 52aaffLuc, 86aapPL-M and 86aapPL were incubated with SRP. Increasing amounts of purified NAC were added to equal aliquots. Subsequent irradiation resulted in crosslinking. **c**, Quantification of the competition between NAC and SRP. The fluorograph of a similar experiment to **b**, but including other nascent chains, was quantified by densitometry. **d**, NAC is the primary cytosolic determinant

of SRP fidelity. For lanes 1–7, salt-stripped 52aaffLuc nascent chains were assembled in the reticulocyte translation system and resuspended in 1/6th vol. TBB and incubated with SRP (20 nM). Then BBC (3 μl per 12-μl assay), NAC-depleted BBC (Heparin F.T. 3 μl per 12 μl), NAC (300 nM) and BSA (5 μM) were added as indicated. The SRP crosslink, indicated by an open circle, persists in the presence of NAC-depleted cytosol and with an excess of BSA, but not when NAC is present. NAC was depleted of 150 μl BBC-A.S. Cut (see Fig. 2a, lane 1) but chromatography on heparin acrylic beads (150 μl dry bed volume, 60 min at 4 °C) and the undiluted flowthrough was used. Lanes 8–11 show Coomassie-blue-stained gels (Coomas.) and immunoblotting (Im.B.) of the starting material (lane 8), the flowthrough (lane 9), and the 400 mM KCl eluate (lane 10). Lane 11 is a NAC standard.

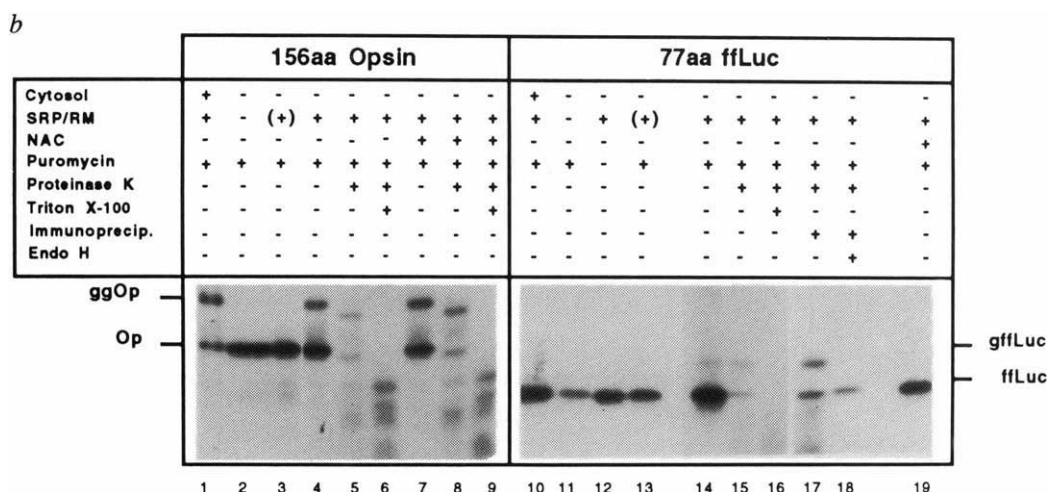
evident (Fig. 5a, lane 1). On addition of membranes containing SRP receptor, the signal peptide was released from SRP¹⁷ and the crosslink was diminished (Fig. 5a, lane 3). The same result was obtained with all nascent chains tested (Fig. 3a, lanes 4–12; and 52aaffLuc, 51aapF¹β and 55aapPL-M, data not shown). Thus, of the nascent chains that could be crosslinked to SRP, approximately equal percentages were targeted to the SRP receptor. Stated differently, in our experimental conditions, SRP alone cannot distinguish between proteins possessing signals from those lacking them, and once bound to any protein, will target the protein to the SRP receptor at the ER membrane.

To test whether inappropriately targeted proteins could be translocated and to analyse the role of NAC in translocation, we used 77aaffLuc and, as a positive control, 156aa opsin nascent chains. Opsin is a multi-membrane-spanning protein in which the first membrane-spanning domain serves as an uncleaved signal peptide. The N-terminal 34 residues, including two *N*-glycosylation sites, are located in the lumen of the ER¹⁸.

The 156aa truncated opsin spans the membrane once and becomes glycosylated *in vitro*¹⁹. Translocation into the lumen of the ER is measured by protease protection of either signal cleaved or *N*-glycosylated proteins²⁰. As 77aaffLuc has no signal peptide, we exploited the fact that a predicted *N*-glycosylation site exists at amino acid 51 which, if used, would allow for detection of translocation.

156aa opsin or 77aaffLuc mRNA was translated in the rabbit reticulocyte system supplemented with SRP and ER membranes. As expected, opsin was transported whereas fLuc was not (Fig. 5b, lanes 1, 10). Surprisingly, when SRP and ER membranes were added post-translationally to salt-stripped 156aa opsin and 77aaffLuc nascent chains, both proteins were transported (Fig. 5b, lanes 4, 14). Transport of the 77aaffLuc depends on puromycin to release the nascent chain from the ribosome. Most of the transported material, but not the untransported protein, remained resistant to proteinase K only in the absence of detergent, showing that glycosylation occurred luminally (Fig. 5b,

FIG. 5 In the absence of NAC, SRP can target 77aaffLuc to the ER membrane and translocation occurs as a consequence. **a**, ER membranes displace SRP from nascent chains. SRP (25 U per 25- μ l assay) bound to salt-stripped nascent chains with or without signal peptides is released when ER membranes (2 equivalents EKRM, for definition and preparation see ref. 29) were added. This is demonstrated by the diminished crosslink of SRP 54K protein to nascent chains in the presence of membranes (compare lanes 1 and 3, 4 and 6, 7 and 9, 10 and 12). **b**, NAC prevents non-signal-containing proteins from being translocated into the ER *in vitro*. SRP (25 U per 25 μ l) was allowed to bind to salt-stripped nascent chains of 156aa bovine opsin and 77aaffLuc which were assembled in the reticulocyte lysate translation system. Evidence for translocation into ER membranes (RM) is given by glycosylation (ggOp, gffLuc; lanes 1, 4, 5, 7, 8 and 14, 15, 17) and by protection against proteinase K (lanes 5, 8, 15, 17). Glycosylation was proven by Endo H digestion (lane 18). Glycosylation failed to occur when puromycin was omitted (lane 12), the RMs were inactivated by NEM (lanes 3, 13) and, in the case of 77aaffLuc, when either cytosol was present (lane 10) or purified NAC (500 nM) was returned to the translocation assay (lane 19). Transport of 156aa opsin occurred in the presence of cytosol (lane 1) or NAC (lanes 7–9). Identity of gffLuc was proven by immunoprecipitation with firefly luciferase antibodies (lanes 17, 18) of five times the amount of material shown in lane 15. Endo H, proteinase K and puromycin treatments and immunoprecipitation are as described (standard protocols as cited in ref. 30).



lanes 5, 6 and 15, 16). The transported 77aaffLuc could be immunoprecipitated with a luciferase antibody, was sensitive to endoglycosidase H (Endo H; Fig. 5b, lanes 17, 18) and bound to concanavalin A-Sepharose (data not shown).

Thus, when SRP inappropriately binds luciferase, it can be translocated, albeit inefficiently. In the presence of cytosol or NAC, opsin, but not luciferase, was transported (Fig. 5b, lanes 7–9, 19). Thus, NAC restores the specificity of transport and prevents mistranslocation. If the membranes were inactivated by *N*-ethylmaleimide (NEM)²¹, neither protein is translocated (Fig. 5b, lanes 3, 13).

Discussion

We have demonstrated that with crosslinking probes incorporated into nascent chains in a variety of amino-acid sequences and at various distances from the peptidyl transferase, NAC crosslinks to all tested nascent chains. We therefore believe that NAC binds ubiquitously to emerging polypeptides. A notable exception is that no crosslink to fully emerged signal peptides could be detected.

The current model of the ribosome maintains that nascent-polypeptide domains traverse the large subunit through a channel that contains about 40 amino acids^{22–25}. NAC is therefore likely to be one of the first non-ribosomal proteins encountered by an emerging polypeptide, because it crosslinks to nascent chains as near as 35 amino acids from the peptidyl transferase centre (see Fig. 1a, lane 7). With 55aapPL, where the modified lysines are no more than 51 amino acids from the peptidyl transferase, supplementation of cytosol with exogenous SRP results in both SRP 54K protein and NAC crosslinks (Fig. 3a, lane 13). Interestingly, cytosol depletion of the nascent chains by salt extraction results in increased crosslinking to SRP 54K protein

(Fig. 3a, lane 14). Readdition of NAC alone restores the NAC crosslink and prevents the SRP crosslink, indicating that for this nascent chain, where we hypothesize that the signal peptide has not fully emerged from the ribosome, NAC and SRP compete for crosslinking. When the signal peptide has fully emerged, as is the case with 86aapPL, SRP crosslinking cannot be competed for by NAC. We can rule out the possibility that NAC fails to compete because the modified lysine is too far from the peptidyl transferase by the fact that 86aapPL-M can crosslink to NAC. In NAC-depleted cytosol, SRP crosslinks nascent chains lacking signal peptides. Readdition of purified NAC restores the specificity of SRP. Although other proteins may be involved in maintaining SRP specificity, NAC is probably responsible for most of the activity (Fig. 4d).

SRP bound to non-signal peptides can mediate targeting to the SRP receptor at the ER membrane. Moreover, SRP is released from the non-signal peptide region (see Fig. 5a). As a consequence, proteins lacking signal peptides that arrive at the ER membrane can be translocated. Purified NAC prevents mistranslocation.

The lower translocation efficiency of non-signal-containing polypeptides is perhaps due to the fact that the signal peptide may serve another function at the membrane such as opening the translocation pore²⁶. We hypothesize that our ribosomes programmed with mRNA for non-secretory proteins, but targeted by SRP, can open the channel, but with decreased efficiency. It is also possible that the proline- and glycine-rich, non- α -helical N terminus of ffLuc is a poor translocation substrate. Interestingly, a SecY mutant strain of *Escherichia coli* was found to be capable of translocating proteins lacking functional signal sequences²⁷, thereby supporting the idea that a signal peptide is not mechanistically required for a protein to cross a membrane.

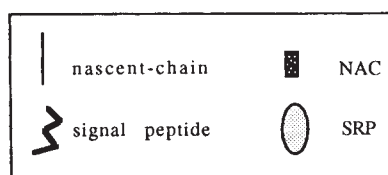
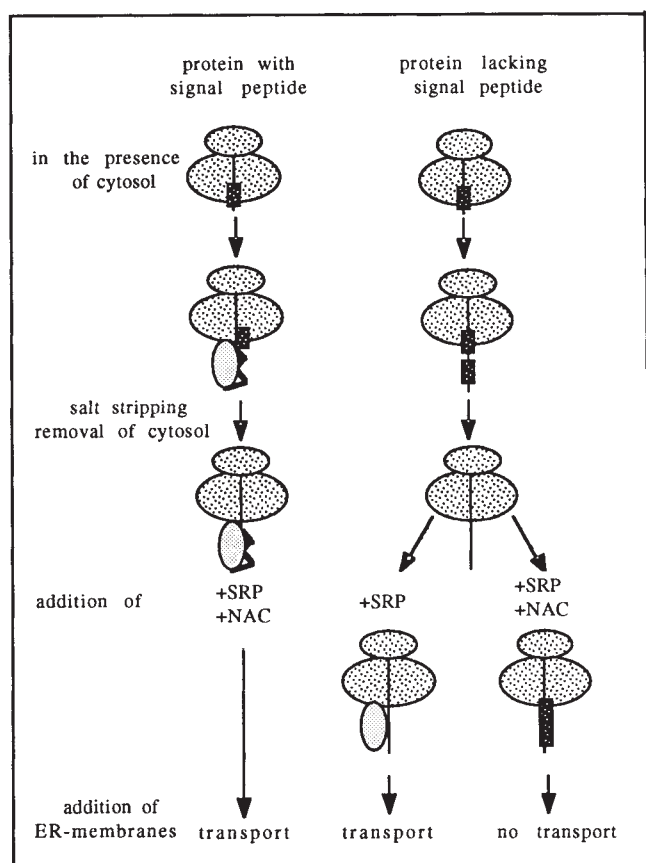


FIG. 6 Summary and current model of NAC function. The ribosome on the left is synthesizing a nascent chain with a signal peptide whereas the one on the right has a nascent chain lacking a signal peptide. If the nascent chain bearing a signal peptide is very short, we observe crosslinking to NAC and not SRP. Salt extraction followed by sedimentation of ribosome-associated nascent chains removes NAC (see Fig. 1b) but does not remove SRP from the signal peptide (see Fig. 3b, lanes 2, 3). After removal of cytosolic factors, SRP can bind nascent chains lacking signal peptides which can, as a result, be translocated into the ER lumen. By competing with SRP for binding to non-signal peptide regions, NAC ensures fidelity in sorting and transport. We have determined that at least two NACs can bind to a single nascent chain (Wang, S., Sakai, H. and Wiedmann, M., unpublished results).

Our working model (see Fig. 6) is that NAC can bind all sequences as they emerge from the ribosome, including partially exposed signal peptides, but is incapable of binding fully exposed signal peptides. In this way, NAC introduces an additional layer of specificity in sorting proteins containing signals from those lacking them. A consequence of our model is that SRP binds signal peptides at least partially by default because they are not

substrates for NAC. Both SRP and NAC have complementary specificities, NAC for non-signal peptides and SRP for signal peptides, but both are necessary for SRP specifically to bind signals. A prediction of our model is that NAC can bind to non-signal-peptide regions of proteins containing signal peptides as shown in Fig. 6. We have proved this by using bifunctional crosslinking reagents (not shown). The NAC interaction with non-signal-peptide regions of 86aapPL is not seen with our photocrosslinking approach because the only modified lysines are in the signal peptide. An additional consequence of SRP binding to non-signal peptides is that these proteins can be translocated, though with low efficiency, into the lumen of the ER.

We hypothesize that NAC may be a component of a dynamic, flexible and partially salt-extractable exit site from the ribosome. We further speculate that NAC in a broader sense functions as an adaptor between ribosomes and the cellular folding and transport machineries and thereby protects nascent chains from premature and inappropriate interactions with them. □

Received 8 April; accepted 4 July 1994.

1. Blobel, G. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1496–1500 (1980).
2. Rapoport, T. A. *Science* **258**, 931–936 (1992).
3. Gething, M.-J. & Sambrook, J. *Nature* **355**, 33–45 (1992).
4. Hendrick, J. P. & Hartl, F.-U. A. *Rev. Biochem.* **62**, 349–384 (1993).
5. Walter, P., Gilmore, R. & Blobel, G. *Cell* **38**, 5–8 (1984).
6. Gilmore, R., Collins, P., Johnson, J., Kellaris, K. & Rapiejko, P. *Meth. Cell Biol.* **34**, 223–239 (1991).
7. Brunner, J. A. *Rev. Biochem.* **62**, 483–514 (1993).
8. Görlich, D., Kurzchalia, T. V., Wiedmann, M. & Rapoport, T. A. *Meth. Cell Biol.* **34**, 241–262 (1991).
9. Lührink, J. et al. *Nature* **359**, 741–743 (1992).
10. Okubo, K. et al. *Nature Genet.* **2**, 173–179 (1992).
11. Kanno, M., Chalut, C. & Egly, J.-M. *Gene* **117**, 219–228 (1992).
12. Walter, P., Ibrahim, I. & Blobel, G. *J. Cell Biol.* **91**, 545–550 (1981).
13. Krieg, U., Walter, P. & Johnson, A. E. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8604–8608 (1986).
14. Kurzchalia, T. V. et al. *Nature* **320**, 634–636 (1986).
15. Walter, P. & Blobel, G. *J. Cell Biol.* **97**, 1693–1699 (1983).
16. High, S., Flint, N. & Dobberstein, B. *J. Cell Biol.* **113**, 25–34 (1991).
17. Wiedmann, M., Kurzchalia, T. V., Bielka, H. & Rapoport, T. A. *J. Cell Biol.* **104**, 201–208 (1987).

18. Friedlander, M. & Blobel, G. *Nature* **318**, 338–343 (1985).
19. Rapiejko, P. J. & Gilmore, R. *J. Cell Biol.* **117**, 493–503 (1992).
20. Blobel, G. & Dobberstein, B. *J. Cell Biol.* **67**, 852–862 (1975).
21. Gilmore, R., Blobel, G. & Walter, P. *J. Cell Biol.* **95**, 463–469 (1982).
22. Malkin, L. I. & Rich, A. *J. molec. Biol.* **26**, 329–346 (1967).
23. Blobel, G. & Sabatini, D. D. *J. Cell Biol.* **45**, 130–145 (1970).
24. Bernabeu, C. & Lake, J. A. *Proc. natn. Acad. Sci. U.S.A.* **79**, 3111–3115 (1982).
25. Yonath, A., Leonard, K. R. & Wittmann, H. G. *Science* **236**, 813–816 (1987).
26. Simon, S. M. & Blobel, G. *Cell* **69**, 677–684 (1992).
27. Derman, A. I., Puziss, J. W., Bassford, P. J. & Beckwith, J. *EMBO J.* **12**, 879–888 (1993).
28. Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 682–691 (1983).
29. Walter, P. & Blobel, G. *Meth. Enzym.* **96**, 84–93 (1983).
30. Wiedmann, M., Kurzchalia, T. V., Hartmann, E. & Rapoport, T. A. *Nature* **328**, 830–833 (1987).

ACKNOWLEDGEMENTS. We thank B. Lauring for discussion and assistance in writing the manuscript; B. Dobberstein, R. Gilmore, F.-U. Hartl, W. Neupert and G. Schatz for plasmids and antibodies; J. Brunner and A. Girschovitch for TDBA; P. Tempst, S. Geromanos and H. Erdjument-Bromage for peptide sequencing; and B. Dobberstein, C. Nicchitta, T. Rapoport, J. Rothman, P. Walter and members of our department for constructive comments. B.W. and H.S. contributed equally to this work.

A faint luminous halo that may trace the dark matter around spiral galaxy NGC5907

Penny D. Sackett*, Heather L. Morrison†, Paul Harding§ & Todd A. Boroson†

* Institute for Advanced Study, Princeton, New Jersey 08540, USA

† US Gemini Project Office, ‡ National Optical Astronomy Observatories, Tucson, Arizona 85726, USA

§ Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA

THE presence of unseen haloes of 'dark matter' has long been inferred from the high rotation speeds of gas and stars in the outer parts of spiral galaxies¹. The volume density of this dark matter decreases less quickly from the galactic centre than does that of the luminous mass (such as that in stars), meaning that the dark matter dominates the mass far from the centre^{1,2}. While searching for faint starlight away from the plane of the edge-on spiral galaxy NGC5907 (ref. 3), we have found that the galaxy is surrounded by a faint luminous halo. The intensity of light from this halo falls less steeply than any known luminous component of spiral galaxies, but is consistent with the distribution of dark mass inferred from the galaxy's rotation curve.

In order to characterize faint extended stellar emission around galaxies it is necessary to (1) calibrate precisely and accurately the detector response, (2) obtain accurate measurements of the sky brightness at distances that avoid contamination from galaxy light, and (3) have many spatial-resolution elements across the galaxy. These requirements motivate the use of a large-format CCD (charge-coupled device) on a wide-field telescope to study large, edge-on systems. We have carried out extremely deep photometry of the thin, edge-on (inclination $i = 87^\circ$) Sc spiral galaxy NGC5907 using a $2,048 \times 2,048$ CCD on the Kitt Peak National Observatory 0.9-m telescope over a field size of ~ 24 arcmin square. By obtaining six 1,800-s galaxy exposures, in which the image was displaced ~ 30 pixels (corresponding to ~ 3 disk scale heights), and 15 dark sky flats of 1,800-s exposure each, we have been able to characterize the response of the detector (flat-field) to 0.02% of sky in 10×10 pixel (8×8 arcsec) bins, and to 0.1% over the large scales (~ 100 arcsec or ~ 5 kpc) typical of the extended halo light. (This high level of flat-fielding accuracy has been achieved in previous work on faint galaxies^{4,5}.) The sky level of $R = 20.43$ mag arcsec⁻² has been determined to 0.01% from the edges of the galaxy frames and subtracted. An accurate error model including the effects of photon noise, flat-fielding errors, readout noise and surface-brightness fluctuations allows us to reach $R = 27$ mag arcsec⁻² reliably. Full details concerning

the data acquisition, reduction, and error analysis are reported elsewhere³.

A spiral galaxy is composed of a thin disk of young stars (called Population I stars) whose local surface brightness falls exponentially with cylindrical distance from the galactic centre and with height above the galactic plane. In addition, one or more of the following stellar components is generally present⁶: (1) a thick disk with larger scale height, (2) a centrally condensed bulge confined to the inner regions of the galaxy, (3) an extended spheroidal stellar halo composed of old (Population II) stars found as field stars or in globular clusters. (The precise relationship between bulges and Population II haloes remains unclear.) In addition to differences in kinematics, these components are distinguished by differences in spatial scales and surface brightnesses: the thin disk and bulge are brighter and less extended than the thick disk and stellar halo. Known Population II field stars and globular cluster systems have luminosity volume densities that fall steeply with distance (r) from the galactic centre: Milky Way globular clusters have an $r^{-3.5}$ density distributions⁷, whereas the M31 halo field star distribution is r^{-4} or steeper⁸; similar results are seen for NGC4565 (ref. 9).

An original goal of this programme was to search for a thick disk in NGC5907. No thick disk was detected, but extended light (in excess of the thin disk) was seen below $R \approx 25$ mag arcsec⁻² in all profiles parallel to the minor axis at distances corresponding to 7–16 thin-disk scale heights (3–7 kpc) above the galactic plane (Fig. 1). Because the surface brightness of this extended light is typically 1% of sky over these regions, neither an inappropriate choice of sky level nor flat-fielding errors can account for its presence. Furthermore, the faint halo is not due to scattered light; by using bright, uncrowded stars (some saturated) to derive a point spread function out to 3 arcmin (10 kpc), we have deconvolved the image and see no change in the extended light³. The extremely small bulge observed in the infrared¹⁰ is too centrally concentrated to account for the excess light, and detailed modelling has indicated that the extended light is not due to a thick disk³, as a scale height ~ 3 times that of any known thick disk (and a normalization at least an order of magnitude less) would be needed.

We therefore model the halo light of NGC5907 with a flattened, power-law distribution with a core, applying a nonlinear least-squares routine to fit thin disk plus power-law halo models to the total light. (Inside the core, the model volume density is roughly constant; outside the core the density drops as an inverse power of the radius.) Our least-squares routine is capable of determining the power-law exponent to within 0.20, and the density axis ratio c/a (that is, the flattening) to within 0.15; these estimates were obtained in experiments using artificial data, in which the power-law behaviour of the profile is known, in the

FIG. 1 The total surface brightness (in the R band) as a function of height above the galactic plane of NGC5907 for four different cuts (spaced 4.1 kpc apart) perpendicular to the galactic plane. Positions of the cuts are shown in Fig. 2a. For each profile, light from matching positions in the four quadrants is averaged. Error bars include differences between quadrants as well as individual errors. The heights above the galactic plane assume a distance of 11 Mpc to the galaxy. Open symbols denote areas underneath the mask shown in Fig. 2, which do not contribute to the halo fit; closed symbols indicate unmasked areas. The dashed line shows the best fit to the total projected light using a thin, exponential disk, with scale length 4.8 kpc, scale height 430 pc, central surface brightness of 19.2 mag arcsec⁻² (in the R band) and truncation radius of 19.5 kpc. To obtain the thin-disk fit, the inner regions of the galaxy were masked to avoid contamination from dust³. Also shown are the best fit for thin-disk plus $r^{-3.5}$ halo with a 200-pc core (dotted line), and thin-disk plus $r^{-2.26}$ halo with a 2-kpc core (solid line).

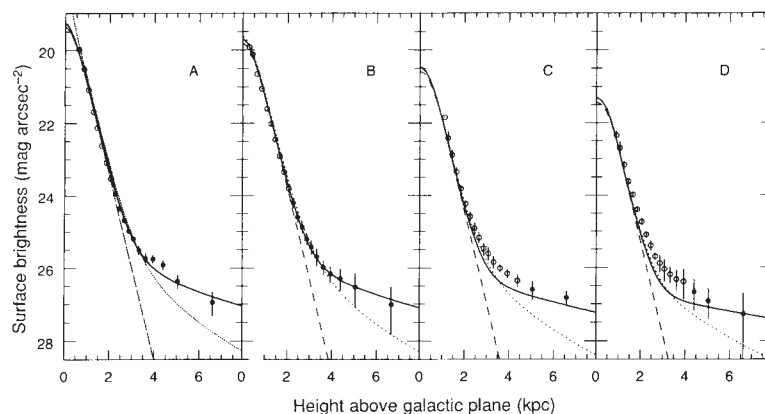
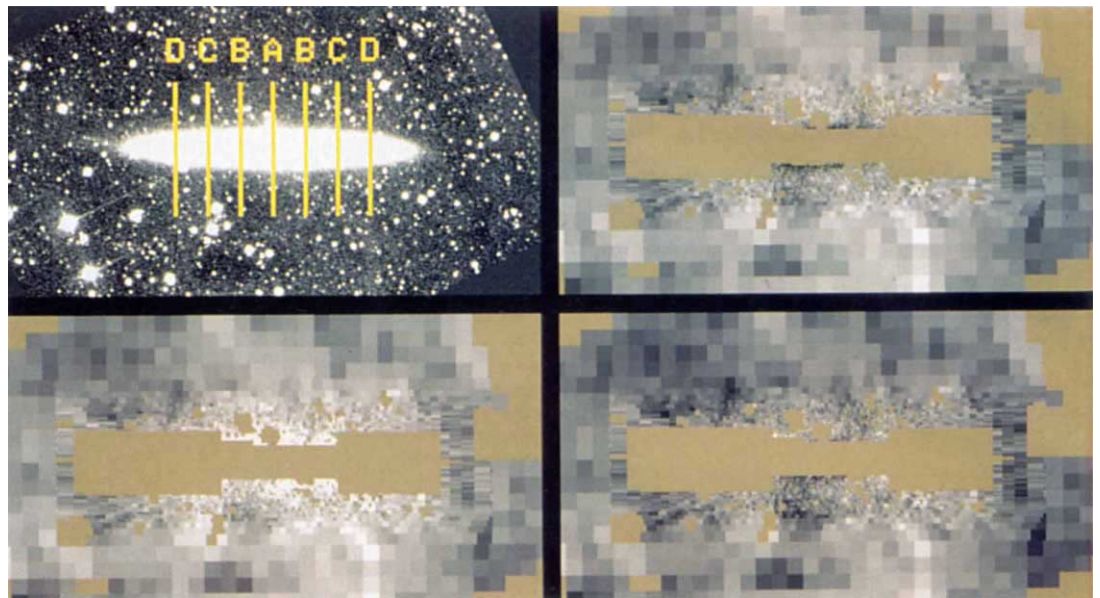


FIG. 2 Optical image of NGC5907 (top left) and weighted residual images for three models. The area shown in each frame is slightly less than one-half the total usable detector area. Light areas indicate regions of excess light compared to the given model; dark areas indicate regions of insufficient light; khaki shows regions that have been masked due to confusion from dust, foreground stars and warp. The data have been binned into 4×4 arcsec bins near the galaxy, and 40×40 arcsec bins in the halo and sky regions, in order to achieve similar signal-to-noise in each bin; thus each bin, regardless of its size, receives a similar weight in the fit. The residuals are expressed in terms of the uncertainty in each bin; white indicates regions with light in excess of 4σ from the model, and black indicates 4σ deficits of light. The masked region is 46 kpc long and 4.3 kpc wide at the minor axis. Top left, grey-scale R-band image with the placement of the vertical profiles shown in Fig. 1 superimposed. The contrast has been adjusted so that the galaxy is 'over-exposed'; essentially all light from the exponential disk is shown as pure white. Bottom left, residuals from thin-disk only model. Disk parameters are



given in Fig. 1 legend. The strong positive residuals indicates the presence of extended halo light. Reduced χ^2 for this fit is 6.0. Top right, residuals from thin-disk plus $r^{-3.5}$ halo model. Disk parameters are given in Fig. 1 legend; halo has $c/a = 0.27$ and a core radius of 200 pc. Reduced χ^2 for this fit is 2.3. Bottom right, residuals from thin disk plus free power-law halo model. Disk parameters are unchanged; halo has an exponent of -2.26 , $c/a = 0.53$, and a core radius of 2 kpc. Reduced χ^2 for this fit is 1.5.

presence of realistic noise. The resulting weighted residuals after the model has been subtracted from the photometric data of NGC5907 are shown in Fig. 2. When the halo exponent is allowed to float as a free parameter, the best fits to the extended light with small cores (<1 kpc, characteristic of known stellar haloes) have an exponent of ~ -2.2 and an axis ratio of ~ 0.5 ; models with a larger core radius of 6 kpc have an exponent of ~ -2.7 and a similar axis ratio.

The surface brightness that we observe for the extended light is appropriate for a stellar halo of Population II field stars¹¹, but its radial profile is much shallower than the $r^{-3.5}$ profile typical of Population II field stars and globular clusters. (NGC5907 has been searched for globular clusters¹² but none were detected.) To our knowledge, this is the first detection of an extended luminous component measured to have an approximately r^{-2} profile around a spiral galaxy. Only in the outer regions of cD galaxies, which have a totally different environment, has a similarly shallow radial profile been observed¹³. The shallowness of its profile suggests that the faint halo around NGC5907 has suffered little from the dynamical evolution processes such as dissipation and relaxation that produce the central concentration and steep profiles of typical stellar haloes¹⁴.

Since the faint extended light of NGC5907 is not well fit by the steep profile of a typical Population II halo (Figs 1, 2), we are led to consider how it may be related to the only component of spirals known to have a shallow distribution: the unseen 'dark haloes' that are inferred to have intrinsic r^{-2} mass profiles. (Somewhat steeper profiles have also been suggested¹⁵.) Dark haloes are generally thought to have core radii of 1–10 kpc (refs 16, 17) and to be significantly flattened¹⁸, but the nature of their constituents remains unknown. The speculative hypothesis that the faint light that we have detected traces (with constant mass-to-light ratio M/L_R) the dark matter halo of NGC5907 predicts that a simple rescaling of the halo light would produce, in combination with the other luminous components of the galaxy, a

model capable of reproducing the observed kinematics. In Fig. 3 we show that a remarkably good match to the rotation curve of NGC5907 (ref. 2) can be obtained using a mass model containing only a (stellar + gaseous) disk and a massive halo with shape and scale parameters derived from the fits to the observed light extrapolated over the radial range of the kinematical data.

Could the faint halo light of NGC5907 be composed of dim, low-mass stars that not only trace the dark matter, but are also responsible for all of its dark mass? By varying only two parameters in the rotation curve fit, the M/L_R of the disk and halo, it is possible to place constraints on the density normalization of the halo required to achieve a good fit; these constraints require that $270 \lesssim M/L_{R, \text{halo}} \lesssim 540$ in order to explain simultaneously the faint-extended light around NGC5907 and its dynamics. Recent reports of microlensing in the Milky Way^{19, 21} suggest that a measurable fraction of the dark matter in our own Galaxy may be in the form of compact objects of roughly one-tenth the mass of the Sun, near the hydrogen burning limit for stars. The magnitudes and colours of M dwarfs and brown dwarfs, which have masses in this range, are strong functions of their metallicity^{22, 23, 24}. If such objects make up the dark halo, they would be expected to have heavy-element abundances at least as small as those of Population II stars. Such stars are seen in the solar neighbourhood, namely the late M subdwarfs of the Galaxy's Population II halo²². The average $M/L_R \approx 450$ that we derive for the putative 'dark' constituents of NGC5907 is slightly lower than the $M/L_R \approx 600$ of faint observed late M subdwarfs²⁴, but higher than the $M/L_R \approx 90$ derived from models of zero-metallicity M dwarfs near the hydrogen-burning limit²³. Detection limits are strongly dependent on wavelength; non-detections of extended light in NGC5907 from older photometry in the B (ref. 25) and K (refs 10, 26) bands are consistent with the colours of observed late M dwarfs. Models of zero-metallicity brown dwarfs give $M/L_R \approx 30,000$ (ref. 24); brown dwarfs are thus too dim to account for the halo light of NGC5907.

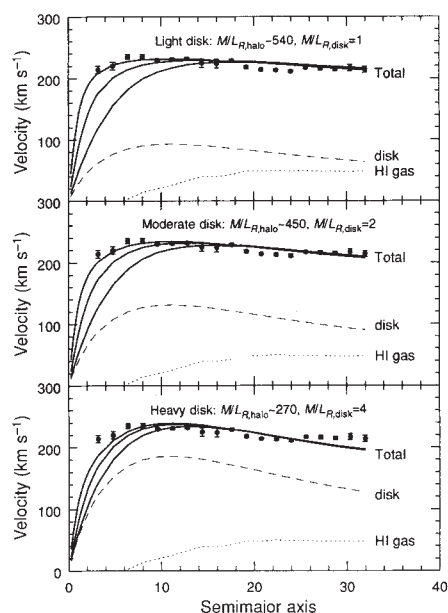


FIG. 3 The rotation curve of the galaxy NGC5907 as derived from observations of the 21-cm line of neutral hydrogen² is shown with model curves. In each panel, the dotted line shows the curve that would be expected from the H I gas alone and the dashed line indicates the rotation curve of the thin exponential disk alone. The thick solid lines are the complete rotation curves, namely the quadrature sum of the gas curve, disk curve, and the rotation curve of haloes producing similarly good fits to the light: an $r^{-2.22}$ halo, with $c/a=0.51$ and a 1-kpc core (top curve); an $r^{-2.26}$ halo, with $c/a=0.53$ and a 2-kpc core (middle curve; R-band residuals shown in Fig. 2d); and an $r^{-2.74}$ halo, with $c/a=0.53$ and a 6-kpc core (bottom curve). The central dark mass densities of these models are 1.9, 0.45 and 0.08 $M_{\odot} \text{ pc}^{-3}$ respectively. In the middle panel, the mass of the disk is assumed to be $5.3 \times 10^{10} M_{\odot}$, derived using $M/L_R=2$ (appropriate to Sc galaxies) and a luminosity based on the Tully-Fisher relation. The M/L_R of the halo component has been scaled to give a good fit to the rotation curve: $M/L_R=420$ for the $r^{-2.22}$ halo, $M/L_R=450$ for the $r^{-2.26}$ halo, and $M/L_R=455$ for the $r^{-2.74}$ halo. Plausible fits to the curve can be obtained with plausible $1 \leq M/L_{R,disk} \leq 4$ (top, middle and bottom panels) for an Sc disk, resulting in $270 \leq M/L_{R,halo} \leq 540$ for the halo, though the high $M/L_{R,disk}$ and low $M/L_{R,halo}$ give notably poorer fits. (All M/L are given in solar units.) The $r^{-3.5}$ distribution cannot provide the required rotation support beyond the edge of the stellar disk (~ 20 kpc). The rotation curve inward of 5 kpc is affected by light inside our mask and by the small bulge component seen in the infrared¹⁰.

As halo stars have high space motions, proper-motion surveys provide one of the best constraints on the existence of a massive halo of M subdwarfs in our own Milky Way. Assuming kinematics similar to the Galaxy's Population II halo, the numbers of M subdwarfs in one such proper-motion catalogue²⁷ fall short by at least an order of magnitude (C. C. Dahn, J. Liebert, H. C. Harris, P. C. Boeshaar, manuscript in preparation), suggesting that the Galaxy's massive halo is not composed of such objects. Deep pencil-beam surveys of the Milky Way capable of accurate star-galaxy separation at $V=26$ mag or fainter would also place tight constraints on a dark-matter halo of late M subdwarfs in the Galaxy, because such surveys would probe distances up to 2 kpc and thus detect ~ 20 of these faint halo stars per square arcmin.

For external galaxies, the decisive observations, which are in reach of current technology, will be surface-brightness measurements in the V and I bands. If the stellar population emitting the faint halo light of NGC5907 is similar to that of known population II stars, its $(V-I)$ colour should be ~ 1.0 (ref. 28); observed late M dwarfs have $(V-I)$ colours near 2.5 (ref. 22). □

Received 22 April; accepted 28 June 1994.

1. Ashman, K. M. *Proc. astr. Soc. Pacif.* **104**, 1109–1138 (1992).
2. Sancisi, R. & van Albada, T. S. in *Int. Astr. Un. Symp. No. 117* (eds Kormendy, J. & Knapp, G. R.) 67–81 (Reidel, Dordrecht, 1987).
3. Morrison, H. L., Boroson, T. A. & Harding, P. *Astr. J.* (in the press).
4. Tyson, J. A. *Astr. J.* **96**, 1–23 (1988).
5. Caldwell, N., Armandoff, T. E., Seitzer, P. & Da Costa, G. D. *Astr. J.* **103**, 840–850 (1992).
6. Freeman, K. C. A. *Rev. Astr. Astrophys.* **25**, 603–632 (1987).
7. Zinn, R. *Astrophys. J.* **293**, 424–444 (1985).
8. Pritchett, C. J. & van den Burgh, S. *Astr. J.* **107**, 1730–1736 (1994).
9. Kormendy, J. & Bruzual, A. G. *Astrophys. J.* **223**, L63–L66 (1978).
10. Barnaby, D. & Thronson, H. A. *Astr. J.* **107**, 1717–1729 (1994).
11. Morrison, H. L. *Astr. J.* **106**, 578–590 (1993).
12. Harris, H. G., Bothun, G. D. & Hesser, J. E. in *Int. Astr. Un. Symp. No. 126* (eds Grindlay, J. E. & Davis Philip, A. G.) 613–614 (Reidel, Dordrecht, 1988).
13. Bahcall, N. A. A. *Rev. Astr. Astrophys.* **15**, 505–540 (1977).
14. Gott, J. R. A. *Rev. Astr. Astrophys.* **15**, 235–266 (1977).
15. Lake, G. & Feinswog, L. *Astr. J.* **98**, 166–179 (1989).
16. Broeils, A. thesis, Univ. Groningen (1992).
17. Buchhorn, M. thesis, Australian natn. Univ. (1992).
18. Sackett, P. D., Rix, H.-W., Jarvis, B. J. & Freeman, K. C. *Astrophys. J.* (in the press).
19. Alcock, C. et al. *Nature* **365**, 621–623 (1993).
20. Aubourg, E. et al. *Nature* **365**, 623–625 (1993).
21. Udalski, A. et al. *Acta astr.* **43**, 289–294 (1993).
22. Monet, D. G. et al. *Astr. J.* **103**, 638–665 (1992).
23. Saumon, D., Bergeron, P., Lunine, J. I., Hubbard, W. B. & Burrows, A. *Astrophys. J.* **424**, 333–344 (1994).
24. Burrows, A., Hubbard, W. B. & Lunine, J. I. in *Proc. 8th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun* (Publ. astr. Soc. Pacif.) (in the press).
25. van der Kruit in *Int. Astr. Un. Symp. No. 117* (eds Kormendy, J. & Knapp, G. R.) 415–424 (Reidel, Dordrecht, 1987).
26. Skrutskie, M. F., Shure, M. A. & Beckwith, S. *Astrophys. J.* **299**, 303–311 (1985).
27. Luyten, W. J. *Publ. Catalogue* (Univ. Minnesota Press, Minneapolis, 1979).
28. Reed, C. B. *Publ. astr. Soc. Pacif.* **97**, 120–125 (1985).

ACKNOWLEDGEMENTS. We are grateful to R. Sancisi for supplying the H I rotation curve of NGC5907, and A. Burrows and J. Liebert for illuminating discussions about low-mass stars. We thank H.-W. Rix for acting as scientific matchmaker and K. Freeman and L. Searle for suggestions that improved the presentation of this paper. P.D.S. is a J. Seward Johnson Fellow at the Institute for Advanced Study; H.L.M. is a Hubble Fellow at the National Optical Astronomy Observatory.

Interstellar oxide grains from the Tieschitz ordinary chondrite

L. R. Nittler, C. M. O'D Alexander, X. Gao, R. M. Walker & E. K. Zinner

McDonnell Center for the Space Sciences and Department of Physics, Washington University, One Brookings Drive, St. Louis, Missouri, 63130-4899, USA

MOST material in the Solar System has an isotopic composition that represents an average of the different stars that contributed material to the protostellar cloud. Primitive meteorites, on the other hand, preserve grains that retain the isotopic signatures of their individual stellar sources¹ and thus provide valuable insight into stellar and galactic evolution, nucleosynthesis, and solar nebular processes. A large number of pre-solar silicon carbide, graphite and diamond grains have now been isolated^{1,2}, but only three interstellar oxide grains have hitherto been recovered^{3–7}, even though oxygen-rich stars are believed to be the dominant source of dust in the Galaxy^{8,9}. We report here the isolation of 21 interstellar oxide grains from the Tieschitz meteorite. The grains exhibit a wide range of oxygen isotope compositions, indicating that they originated in several distinct stellar sources having different masses and initial compositions. There is also evidence for the presence of the short-lived radionuclide ²⁶Al in nine of the grains at the time they formed. Although the isotopic compositions of many of the grains are consistent with both observations and theoretical models of oxygen-rich red giant stars, a significant fraction have no observed stellar counterpart.

Different isotopes are produced by a variety of nucleosynthetic processes, and the isotopic ratios in interstellar grains can therefore provide insight into the types of stellar sites that produced them. Here we determine the isotopic compositions of 21 inter-

stellar oxide grains isolated from the Tieschitz meteorite, and compare these values to astronomical observations and theoretical models of stellar evolution.

A sample of the Tieschitz (H/L 3.6) ordinary chondrite was physically and chemically processed to produce a residue^{10,11}, T8, in which chemically-resistant oxide phases were concentrated by a factor of $\sim 2.5 \times 10^5$. For ion microprobe analysis, a suspension of T8 was deposited on a gold foil¹² along with grains of the Burma Spinel oxygen isotope standard. To automatically locate rare interstellar oxides, a Photometrics CCD (charge-coupled device) camera was coupled to the microchannel plate/fluorescent screen of the modified Cameca IMS-3F ion microprobe⁴ at Washington University. Low-mass-resolution ion images (in $^{16}\text{O}^-$ and $^{18}\text{O}^-$) of oxide grains were digitized, and the $^{16}\text{O}/^{18}\text{O}$ ratios ($\sigma \approx 4\%$) of individual grains were determined by image processing. Grains that deviated in duplicate analyses by more than 3σ from the solar $^{16}\text{O}/^{18}\text{O}$ ratio were selected for high-mass-resolution analysis. The high mass-resolving power needed to separate $^{16}\text{OH}^-$ ions from the $^{17}\text{O}^-$ ions precludes measurement of $^{16}\text{O}/^{17}\text{O}$ ratios by ion imaging in our instrument.

Ion imaging of $\sim 6,000$ grains yielded 53 candidates, the $^{16}\text{O}/^{18}\text{O}$ and $^{16}\text{O}/^{17}\text{O}$ ratios of which were subsequently measured at high mass resolution. Of these grains, 20 corundum (Al_2O_3) and 1 spinel (MgAl_2O_4) have a large range of anomalous (non-solar) $^{16}\text{O}/^{18}\text{O}$ and $^{16}\text{O}/^{17}\text{O}$ ratios (Fig. 1 and Table 1). Nine of the grains also have large ^{26}Mg excesses, corresponding to initial $^{26}\text{Al}/^{27}\text{Al}$ ratios of between 1.2×10^{-4} and 7.8×10^{-3} , much higher than the maximum value of 5×10^{-5} observed in primitive Solar System material (Fig. 2 and Table 1). All grains are 0.5–2 μm in size. For purposes of discussion, the 24 interstellar oxide grains found to date were divided into three groups on the basis of their oxygen isotope compositions (Fig. 1 and Table 1). Note that the range of oxygen isotope ratios usually observed in meteorites and terrestrial samples falls within the solar symbol in Fig. 1.

Group 1 grains have significant enrichments in ^{17}O and modest depletions in ^{18}O , similar to the isotopic compositions measured in O-rich red giants^{13,14} (Table 1 and Fig. 1). These stars are thought to produce 65–75% of all dust, or ~ 80 –90% of O-rich dust, in the Galaxy^{8,9}. Because of this similarity, Group 1 grains probably formed in the atmospheres of red giants. The oxygen isotope compositions of such stars have been successfully reproduced by theoretical models in terms of the so-called first dredge-up, which occurs after exhaustion of H in the stellar core^{15–18}. In this process, material that has undergone partial core H-burning via the CNO-cycle is mixed into the envelope. These models predict that the first dredge-up significantly decreases the envelope's initial $^{16}\text{O}/^{17}\text{O}$ ratio, the actual values depending primarily on stellar mass^{16–18} (see Fig. 1). But the first dredge-up has only a relatively small effect on the $^{16}\text{O}/^{18}\text{O}$ ratio (20–50% increase), and larger differences in this ratio must be due to differences in the initial isotopic compositions of different stars¹⁷ (Fig. 1). Such differences are the result of Galactic chemical evolution, reflecting age differences of the stars and/or the spread in chemical compositions observed in newly formed stars within a given Galactic epoch¹⁹. The range of the oxygen isotope compositions observed in the Group 1 grains indicates that they originated from several distinct stellar sources with different masses as well as different initial compositions.

As red giant stars continue to evolve, they can undergo two additional dredge-up episodes¹⁵. The second dredge-up occurs only in $>5M_\odot$ stars at the beginning of the asymptotic giant branch (AGB) phase, when He is exhausted in the star's core. The third dredge-up, which is experienced by all stars in the 1–8 $M_\odot$ range, occurs during the thermally pulsing (TP) AGB phase when H and He burn alternately in thin shells on top of an inert core consisting now of C and O. These mixing episodes are not expected to significantly change the oxygen isotope composition of the envelope, although some spectroscopic observa-

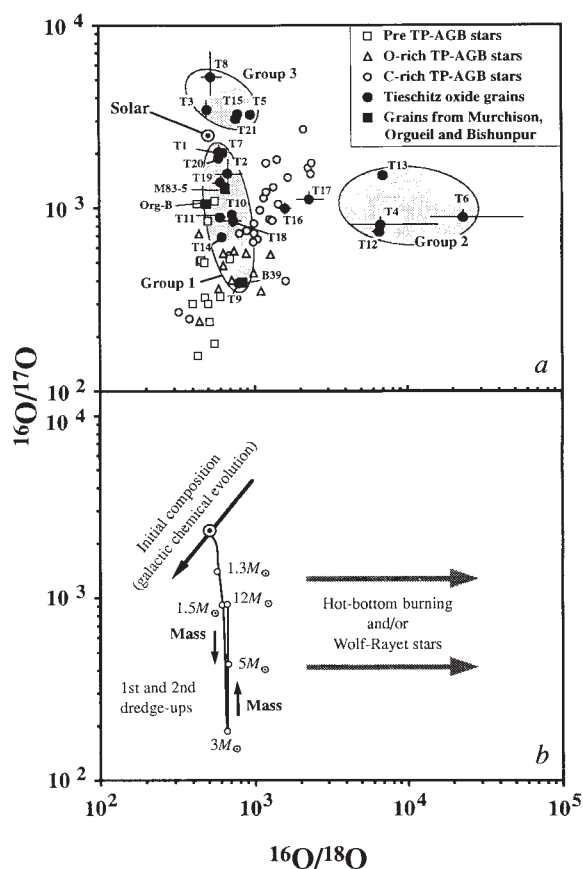


FIG. 1 *a*, Comparison of the oxygen isotope compositions of 24 interstellar oxide grains reported here and elsewhere^{3–7} with those of red giant stars^{13,14,20,21}. The red giant stars in *a* have been divided into three groups based on their spectral types and models of their evolution¹⁵. The oxide grains have also been divided into three groups based on their oxygen isotope compositions. *b*, Diagram of the expected oxygen isotope compositions of red giant envelopes after the first and second dredge-ups as a function of stellar mass and initial composition. After core H-burning ceases, ^{17}O -enriched H-burnt material from the interior is mixed into the O-rich envelope of a red giant during two dredge-up episodes¹⁵ (the second only in $>5M_\odot$ stars). The curve shows the predicted post-dredge-up ratios of stars with initially solar compositions but different masses¹⁸. The initial $^{16}\text{O}/^{17}\text{O}$ ratio influences the final ratio only in (1–1.5) $M_\odot$ stars¹⁷. The $^{16}\text{O}/^{18}\text{O}$ ratio is only slightly affected by these two dredge-ups¹⁷. The range of $^{16}\text{O}/^{17}\text{O}$ and $^{16}\text{O}/^{18}\text{O}$ ratios amongst Group 1 grains suggests that they come from several stars with different masses and different initial compositions. The initial oxygen isotope compositions, reflecting Galactic evolution, probably varied roughly as indicated (F. Timmes, personal communication). Subsequently, during the thermally pulsing asymptotic giant branch (TP-AGB) phase a third multiple dredge-up episode brings ^{12}C -rich He-burnt material to the surface. As a result, (1–5) $M_\odot$ stars will eventually become carbon stars ($\text{C}/\text{O} > 1$), but for (5–8) $M_\odot$ stars, H-burning at the base of the envelope (hot-bottom burning) is predicted to destroy ^{12}C and ^{18}O , thereby preventing formation of a carbon star³¹. Note that oxide grains may form even in carbon-star envelopes³².

tions of TP-AGB stars suggest that their atmospheres have higher $^{16}\text{O}/^{18}\text{O}$ and $^{16}\text{O}/^{17}\text{O}$ ratios than pre-TP-AGB stars^{13,14,20,21} (Fig. 1*a*).

The third dredge-up is expected to bring to the star's surface ^{26}Al that was produced by shell H-burning at much higher temperatures than those reached during core H-burning^{22,23}. Models predict envelope $^{26}\text{Al}/^{27}\text{Al}$ ratios in the range from 5×10^{-4} to 10^{-2} , depending on stellar mass, mass loss rate and evolutionary stage during the TP-AGB phase. Eight of the ten Group 1 grains measured for Al-Mg have inferred initial $^{26}\text{Al}/^{27}\text{Al}$ ratios in this range (Fig. 2). An alternative production mechanism for ^{26}Al is H-burning at the base of the convective envelope of TP-AGB

TABLE 1 Isotopic compositions of 24 meteoritic interstellar oxide grains

Grain	$^{17}\text{O}/^{16}\text{O}$	$^{18}\text{O}/^{16}\text{O}$	$^{26}\text{Mg}^*/^{24}\text{Mg}$	$^{27}\text{Al}/^{24}\text{Mg}$	$^{26}\text{Al}/^{27}\text{Al}$	Group
Solar	3.83×10^{-4}	2.01×10^{-3}				
T1	$5.01 (39) \times 10^{-4}$	$1.66 (12) \times 10^{-3}$	≤ 0.033	209 (32)	$\leq 1.6 \times 10^{-4}$	1
T2	$6.6 (1.1) \times 10^{-4}$	$1.44 (27) \times 10^{-3}$	2.04 (19)	263 (65)	$7.8 (2.0) \times 10^{-3}$	1
T7	$5.05 (22) \times 10^{-4}$	$1.55 (7) \times 10^{-3}$	0.369 (15)	139 (18)	$2.65 (37) \times 10^{-3}$	1
T9	$2.63 (5) \times 10^{-3}$	$1.27 (5) \times 10^{-3}$	0.586 (83)	1086 (98)	$5.40 (90) \times 10^{-4}$	1
T10	$1.11 (3) \times 10^{-3}$	$1.36 (6) \times 10^{-3}$	≤ 0.042	402 (37)	$\leq 1.0 \times 10^{-4}$	1
T11	$1.14 (5) \times 10^{-3}$	$1.60 (10) \times 10^{-3}$	NA	NA	NA	1
T14	$1.47 (5) \times 10^{-3}$	$1.57 (9) \times 10^{-3}$	0.163 (81)	1351 (271)	$1.20 (64) \times 10^{-4}$	1
T18	$1.19 (8) \times 10^{-3}$	$1.35 (15) \times 10^{-3}$	NA	NA	NA	1
T19	$7.50 (74) \times 10^{-4}$	$1.63 (19) \times 10^{-3}$	NA	NA	NA	1
T20	$5.38 (16) \times 10^{-4}$	$1.66 (47) \times 10^{-3}$	0.0303 (14)	41 (2)	$7.34 (46) \times 10^{-4}$	1
M83-5	$7.92 (23) \times 10^{-4}$	$1.52 (5) \times 10^{-3}$	1.96 (11)	2250 (109)	$8.73 (64) \times 10^{-4}$	1
Org-B	$9.72 (10) \times 10^{-4}$	$1.99 (4) \times 10^{-3}$	1.61 (3)	1840 (90)	$8.7 (1) \times 10^{-4}$	1
B39	$2.60 (6) \times 10^{-3}$	$1.17 (4) \times 10^{-3}$	0.2336 (104)	136 (14)	$1.7 (0.2) \times 10^{-3}$	1
T16	$1.01 (6) \times 10^{-3}$	$6.21 (54) \times 10^{-4}$	NA	NA	NA	1/2
T17	$9.12 (76) \times 10^{-4}$	$4.30 (89) \times 10^{-4}$	NA	NA	NA	1/2
T4	$1.26 (11) \times 10^{-3}$	$1.47 (85) \times 10^{-4}$	0.064 (13)	17.0 (2.3)	$3.77 (91) \times 10^{-3}$	2
T6	$1.14 (8) \times 10^{-3}$	$4.3 (2.6) \times 10^{-5}$	0.423 (57)	106 (16)	$4.01 (82) \times 10^{-3}$	2
T12	$1.36 (4) \times 10^{-3}$	$1.50 (11) \times 10^{-4}$	0.969 (81)	135 (11)	$7.19 (83) \times 10^{-3}$	2
T13	$6.72 (24) \times 10^{-4}$	$1.41 (10) \times 10^{-4}$	10.4 (13)	5442 (707)	$1.90 (35) \times 10^{-3}$	2
T3	$2.92 (29) \times 10^{-4}$	$1.98 (13) \times 10^{-3}$	≤ 0.0048	8.2 (1)	$\leq 6.0 \times 10^{-4}$	3
T5	$3.15 (11) \times 10^{-4}$	$1.02 (3) \times 10^{-3}$	≤ 0.0099	47 (6)	$\leq 2.1 \times 10^{-4}$	3
T8	$1.92 (51) \times 10^{-4}$	$1.85 (29) \times 10^{-3}$	NA	NA	NA	3
T15	$3.20 (23) \times 10^{-4}$	$1.26 (6) \times 10^{-3}$	NA	NA	NA	3
T21	$3.25 (17) \times 10^{-4}$	$1.29 (59) \times 10^{-3}$	≤ 0.025	549 (33)	$\leq 4.6 \times 10^{-5}$	3

The (meteorite) source of the grains is shown in the first column; Tieschitz (T), Murchison (M)⁴, Orgueil (Org)^{3,5,6} and Bishunpur (B)⁷. The grains have been divided into three groups on the basis of their oxygen isotopic compositions. $^{26}\text{Mg}^*$ is the excess of ^{26}Mg after correction for mass fractionation, and $^{26}\text{Al}/^{27}\text{Al}$ is the initial ratio if all $^{26}\text{Mg}^*$ results from the *in situ* decay of ^{26}Al . All errors, in parenthesis, are 1σ and upper limits are 2σ . All grains were analysed by SEM-EDS (Scanning electron microscopy—energy dispersive spectroscopy) after oxygen isotope measurements but before Mg-Al analysis in the ion probe. All but grain T3 appear to be corundum (Al_2O_3). The Al/Mg ratio of grain T3, measured by EDS, is about twice that expected for pure spinel (MgAl_2O_4) but much lower than those of the other grains. The ratio obtained in the subsequent ion-probe analysis of T3 is even higher, suggesting that the grain may be an intergrowth of spinel and corundum. The oxygen isotope ratios are given with ^{16}O in the denominator because in the reverse case the errors are asymmetric and nonlinear.

stars of $>5M_\odot$ (hot-bottom burning)^{24,25}. But during hot-bottom burning the whole envelope is cycled through the high-temperature H-burning zone, resulting in the destruction of essentially all ^{18}O (producing high $^{16}\text{O}/^{18}\text{O}$ ratios), in disagreement with the oxygen isotope ratios observed in Group 1 grains. The oxygen isotope and $^{26}\text{Al}/^{27}\text{Al}$ ratios of these grains thus indicate that at least eight of them come from TP-AGB stars, and at least two from red giants before they reached the TP-AGB phase. It is worth noting that whereas the oxygen isotopes in red giant envelopes can be measured astronomically—albeit

with larger errors than those obtained in the grains—the determination of $^{26}\text{Al}/^{27}\text{Al}$ ratios in such stars is only possible from the laboratory analysis of pre-solar grains.

Unlike grains in Group 1, neither Group 2 nor Group 3 grains have spectroscopic counterparts and we must rely solely on comparisons with stellar evolution models to infer their likely sources. Group 2 grains have ^{17}O and ^{26}Al enrichments, and ^{18}O depletions that are much larger than previously observed in any meteoritic material or star. Two other grains, which lie on the edge of the carbon-star field, have intermediate ^{18}O depletions. Relatively low-temperature hot-bottom burning could destroy essentially all ^{18}O in an AGB envelope without significantly changing the ^{26}Al and ^{17}O abundances already established by the first, second and third dredge-ups¹⁷. Also, such isotopic compositions are predicted to exist at the surface of massive mass-losing stars which have shed their envelopes, exposing the H-burnt interior (Wolf-Rayet stars during the Of-WN phases)²⁶. In any case, better modelling and more elemental and isotopic data on Group 2 grains are necessary for distinguishing between the two possible sources.

The Group 3 grains are moderately depleted in ^{17}O relative to the Solar System. They could conceivably have formed around low-mass red giant stars that have experienced the first dredge-up, provided that the initial $^{16}\text{O}/^{17}\text{O}$ ratios of the stars were higher than the measured grain values and, therefore, the solar value. On the other hand, massive stars ($>10 M_\odot$), which contribute ~ 6 – 12% of all Galactic dust^{8,9}, produce large excesses of ^{16}O and ^{18}O in certain shells²⁷ that could come to the surface as the star loses mass or could be ejected in a supernova explosion.

^{16}O enrichments and depletions with respect to the terrestrial isotopic ratios are common in materials of Solar System origin²⁸; in meteorites the enrichments can reach up to 7% in corundum (Al_2O_3) and spinel (MgAl_2O_4) from Ca-Al-rich inclusions²⁹. These enrichments prompted suggestions that ^{16}O -rich grains, probably corundum and spinel, from a supernova were incompletely mixed into the Solar System²⁸. We have not found any

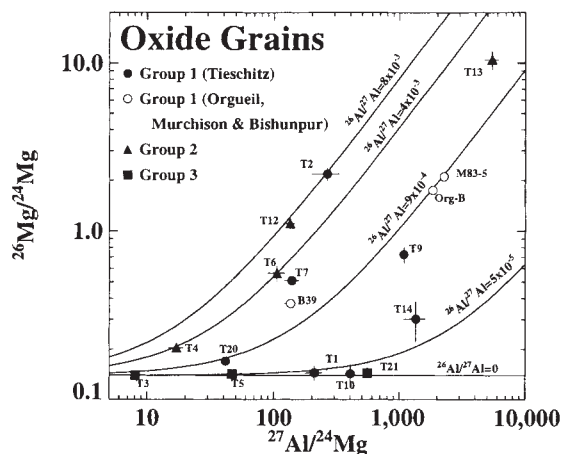


FIG. 2 Plot of the $^{26}\text{Mg}^*/^{24}\text{Mg}$ ratio versus the $^{27}\text{Al}/^{24}\text{Mg}$ ratio in 17 pre-solar oxide grains. The large ^{26}Mg excesses, compared to solar Mg isotopic ratios, are best explained as the result of *in situ* decay of ^{26}Al . Also shown are Al-Mg evolution lines for five initial $^{26}\text{Al}/^{27}\text{Al}$ ratios. The $^{27}\text{Al}/^{24}\text{Mg}$ ratios in most of the Tieschitz corundum grains are not as high as might be expected, possibly because of significant background contributions from isotopically normal Mg to the Mg analysis. However, these contributions do not affect the inferred $^{26}\text{Al}/^{27}\text{Al}$ ratios.

pre-solar grains with large ^{16}O enrichments. The Group 3 grains are ^{16}O -rich, but not with respect to both ^{17}O and ^{18}O . The discovery that ^{16}O can undergo mass-independent gas-phase chemical fractionation from both ^{17}O and ^{18}O has provided an alternative explanation for the observed ^{16}O enrichments in Solar System material³⁰.

Finally, if the Si/Al ratio is assumed to be solar in all stars and if all Al goes into corundum (Al_2O_3), interstellar corundum appears to be underabundant in meteorites, relative to interstellar SiC, by about a factor of between 20 and 50 (refs 4, 6, 10)

when compared to estimated Galactic dust production rates^{8,9}. This is in spite of the fact that, in the solar nebula, corundum should have been more stable than SiC. Possible explanations are that interstellar corundum has a finer grain size distribution than SiC and was thus not detected by our technique, that Al primarily condenses in other phases (such as silicates) that are less resistant to the chemical treatments used to isolate the grains or, most speculatively, that corundum has a shorter lifetime in the interstellar medium than SiC. □

Received 25 February; accepted 28 June 1994.

- Anders, E. & Zinner, E. K. *Meteoritics* **28**, 490–514 (1993).
- Ott, U. *Nature* **364**, 25–33 (1993).
- Huss, G. R., Hutcheon, I. D., Wasserburg, G. J. & Stone, J. *Lunar planet. Sci.* **XXIII**, 563–564 (1992).
- Nittler, L. R., Walker, R. M., Zinner, E. K., Hoppe, P. & Lewis, R. S. *Lunar planet. Sci.* **XXIV**, 1087–1088 (1993).
- Huss, G. R., Hutcheon, I. D., Fahey, A. J. & Wasserburg, G. J. *Meteoritics* **28**, 369–370 (1993).
- Hutcheon, I. D., Huss, G. R., Fahey, A. J. & Wasserburg, G. J. *Astrophys. J.* **425**, L97–L100 (1994).
- Huss, G. R., Fahey, A. J., Gallino, R. & Wasserburg, G. J. *Astrophys. J.* (in the press).
- Gehr, R. D. in *Interstellar Dust* (eds Allamandola, L. J. & Tielens, A. G. G.) 445–453 (Kluwer Academic, Dordrecht, 1989).
- Whittet, D. C. B. *Dust in the Galactic Environment*, 295 (Institute of Physics, Bristol, 1992).
- Nittler, L. R., Alexander, C. M. O'D., Gao, X., Walker, R. M. & Zinner, E. K. *Lunar planet. Sci.* **XXV**, 1005–1006 (1994).
- Gao, X., Alexander, C. M. O'D., Swan, P. D. & Walker, R. M. *Lunar planet. Sci.* **XXV**, 401–402 (1994).
- Zinner, E. K., Tang, M. & Anders, E. *Geochim. cosmochim. Acta* **53**, 3273–3290 (1989).
- Harris, M. J. & Lambert, D. L. *Astrophys. J.* **285**, 674–682 (1984).
- Smith, V. V. & Lambert, D. L. *Astrophys. J. Suppl. Ser.* **72**, 387–416 (1990).
- Iben, I. Jr. *Astrophys. J. Suppl. Ser.* **76**, 55–114 (1991).

- Dearborn, D. S. P. *Phys. Rep.* **210**, 367–382 (1992).
- Boothroyd, A. I., Sackmann, I. J. & Wasserburg, G. J. *Astrophys. J.* (in the press).
- El Eid, M. F. *Astr. Astrophys.* **285**, 915–928 (1994).
- Edvardsson, B. et al. *Astr. Astrophys.* **275**, 101–152 (1993).
- Harris, M. J., Lambert, D. L., Hinkle, K. H., Gustafsson, B. & Eriksson, K. *Astrophys. J.* **316**, 294–304 (1987).
- Kahane, C., Cernicharo, J., Gomez-Gonzalez, J. & Guelin, M. *Astr. Astrophys.* **256**, 235–250 (1992).
- Forestini, M., Paulus, G. & Arnould, M. *Astr. Astrophys.* **252**, 597–604 (1991).
- Gallino, R., Raiteri, C. M., Busso, M. & Matteucci, F. *Astrophys. J.* (in the press).
- Nørgaard, H. *Astrophys. J.* **236**, 895–898 (1980).
- Cameron, A. G. W. in *Protostars and Planets III* (eds Levy, E. H. & Lunine, J. I.) 47–73 (Univ. Arizona Press, 1993).
- Prantzos, N., Doom, C., Arnould, M. & de Loore, C. *Astrophys. J.* **304**, 695–712 (1986).
- Woosley, S. E. in *Nucleosynthesis and Chemical Evolution* (eds Hauck, B., Maeder, A. & Meynet, G.) 1–195 (Observatoire de Genève, Geneva, 1986).
- Clayton, R. N., Grossman, L. & Mayeda, T. K. *Science* **182**, 485–488 (1973).
- Virag, A., Zinner, E. K., Amari, S. & Anders, E. *Geochim. cosmochim. Acta* **55**, 2045–2062 (1991).
- Thiemens, M. H. & Heidenreich, J. E. *Science* **219**, 1073–1075 (1983).
- Renzini, A. & Voli, M. *Astr. Astrophys.* **94**, 175–193 (1981).
- Lattimer, J. M., Schramm, D. N. & Grossman, L. *Astrophys. J.* **219**, 230–249 (1978).

ACKNOWLEDGEMENTS. We thank G. Huss and A. G. W. Cameron for their constructive reviews. This work was supported by NASA.

Directional motion of brownian particles induced by a periodic asymmetric potential

Juliette Rousselet*, Laurence Salome*,
Armand Ajdari† & Jacques Prost†

* Centre de Recherche Paul Pascal, Avenue A. Schweitzer,
33600 Pessac, France

† Groupe de Physico Chimie Théorique, URA 1382 CNRS, E.S.P.C.I.,
10, rue Vauquelin, 75231 Paris Cédex 05, France

STRUCTURES possessing spatial asymmetry should act as pumps in the presence of dissipation alone^{1–4}, without the need for macroscopic forces or temperature differences⁵ to drive vectorial motion. It has been shown theoretically^{2–4,6,7} that particles subjected to an asymmetric periodic potential can display net directional motion even if the space-averaged force is zero. Here we demonstrate such behaviour experimentally. We have studied the behaviour of colloidal particles suspended in solution and exposed to a sawtooth dielectric potential which is turned on and off periodically. The particles exhibit net motion with a velocity that depends on their size, suggesting applications in separation processes for objects in the size range 0.1–5 μm —a range that includes biological structures such as viruses, cells and chromosomes⁸. We furthermore point out the analogy between our device and motor protein assemblies.

Consider an asymmetric potential U_{on} applied periodically for a period τ_{on} and then switched off for a period τ_{off} (Fig. 1A). At the end of an 'on' period, the particles are trapped in the minima of the potential U_{on} , so that the concentration of particles is peaked around the corresponding positions (Fig. 1B). During the following 'off' period the particles diffuse freely so that the concentration at the end of this 'off' period is a set of gaussian curves centred around the same points (Fig. 1C). Turning the potential on again will induce 'downhill' motion of particles (Fig. 1A). The result of this off-on cycle is that the particles

corresponding to the hatched areas move to the right, whereas essentially none move to the left. In a more general picture a proportion $P(m)$ would make an m -step progression, with $P(m) > P(-m)$ owing to the asymmetry. The result is to achieve macroscopic drift in a single direction.

In our experiments, colloidal particles were confined between two glass slides and subjected to a spatially asymmetric and periodic a.c. electric field E , which was successively turned on and off. The field was generated by interdigitated electrode deposited on one of the glass slides with standard photolithographic techniques (Innovations Couches Minces Co., Le Coudray Montceaux, France), the shape of which provided the necessary asymmetry (Fig. 2). The 'Christmas tree' design was chosen in such a way that the dielectric energy profile ($-\frac{1}{2}\Delta\alpha E^2$; where $\Delta\alpha$ is the polarizability of the particle relative to the suspending solution) along a line between two adjacent electrodes, was similar to that in Fig. 1, with a 50- μm period. The neck width between two adjacent electrodes was 5 μm , giving fields as high as $2 \times 10^4 \text{ V cm}^{-1}$ for applied voltages of just 10 V.

Experiments were performed on functionalized polystyrene latex spheres of diameter 0.25 μm , 0.4 μm (fluorescent particles provided by Molecular Probes, Eugene, Oregon) and 1 μm (provided by Interchim, Montluçon, France). The particles were suspended solution (tris (tris(hydroxymethyl)aminomethane) 44.5 mM, boric acid 44.5 mM, EDTA (ethylenediaminetetraacetic acid) 1.2 mM, in water), promoting substantial dissociation of the carboxylate groups at their surface. These conditions were preferred to 'pure' water both because of the easy control of the solution, and its wide use in biological systems. Electrodes were not protected by a passivation layer, and in order to avoid electrolysis a.c. voltages with frequencies higher than 500 Hz were applied. The particle diameters were measured by light-scattering techniques and found to be within 10% of the specifications; their diffusion constants were measured from direct observation of the brownian motion between glass slides, and found to agree with the Stokes-Einstein formula within 15% (diffusion close to the walls is discussed in a preprint by L. Faucheux and A. Libchaber, NEC Research Institute, Princeton NJ). Their dielectrophoretic response was also studied independently, and revealed

a complex behaviour with two relaxation frequencies (J.R., unpublished results). Correspondingly, in our experiment the particles were clearly attracted into the necks of the electrode pattern for two frequency ranges, the second range depending on the size of the particle. At these frequencies, it was easy to achieve conditions such that the trapping energy was significantly larger than kT . This was simply checked by observing that the escape probability was vanishingly small.

A direct analysis of video images allowed us to count the number of particles migrating from one 'trap' to the attraction zone of the next one at each temporal on/off cycle (Fig. 3), and thus to evaluate the average induced velocity. Almost no particles were able to make backward steps. Particles of each size were studied separately at the optimal frequency defined above. Data on 20–50 different traps were accumulated, and several particle densities were investigated. The results reported here correspond to a 'low-density' regime, in that the observed diffusion process during the 'off' period was consistent with single-particle diffusion. The number of particles per trap never exceeded 20 (note however that it allowed the formation of 'strings of pearls' during the 'on' state for the 1- μm particles). Because of the time limitation imposed by the adsorption of particles on the electrodes, we proceeded in two steps for each sample. First, the border between the attraction zones of neighbouring traps was determined (Fig. 2)—its actual shape, intermediate between a circle centred on the electrode neck and a straight line perpendicular to the polar axis, shows that a complete description of the electric-field geometry requires more than a single typical length a . Second, particles were released after trapping, and their diffusion observed. We plot in Fig. 4 the fraction p of particles having crossed the above-mentioned border as a function of the 'off' time τ_{off} . We checked in the case of 0.4- μm particles that this procedure gave the same result as the counting of actual jumps during on/off cycles.

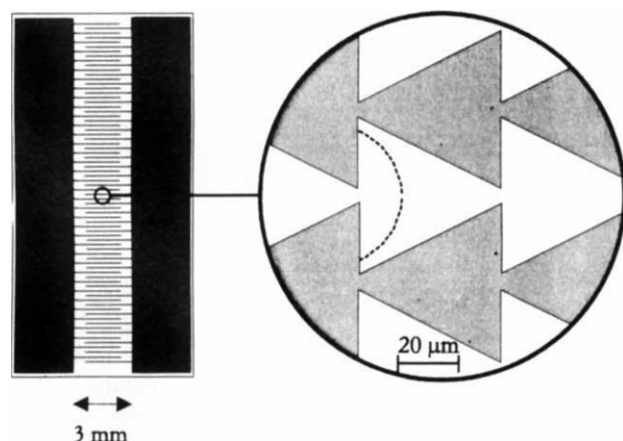


FIG. 2 a, Schematic representation of the interdigitated electrodes. b, Magnified view, and characteristics of the electrical geometry at high frequencies, as inferred from direct observation of particle motion when a voltage difference is applied between the electrodes. Particles are trapped slightly to the right of the electrode necks. The boundary (dotted line) between the attraction zones of neighbouring traps is found to be roughly circular, the minimum distance from the trap to the boundary being a .

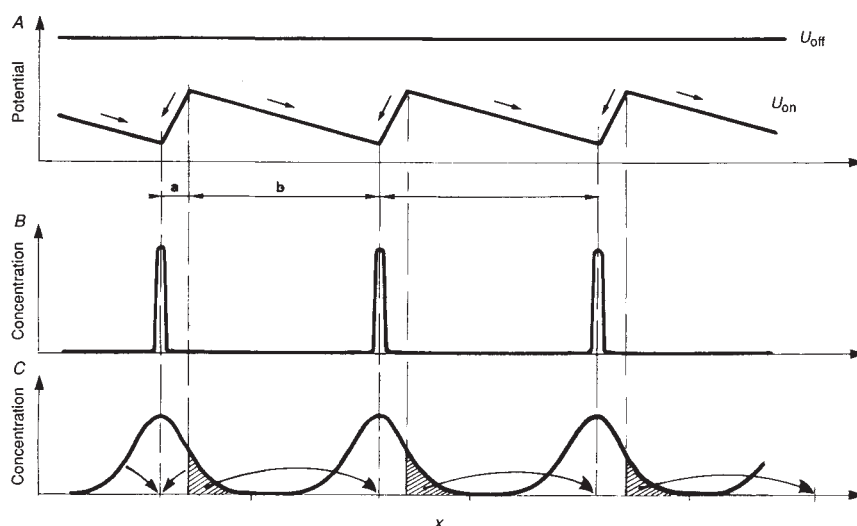


FIG. 1 Schematic illustration of the pumping principles.

In the simplest theoretical picture (isotropic two-dimensional diffusion) p is bounded by the probability $p_1 = 1/2 \exp(-a^2/4D\tau_{\text{off}})$ corresponding to a circular boundary and the probability $p_2 = 1/2 \operatorname{erfc}(a/(4D\tau_{\text{off}})^{1/2})$ corresponding to a straight boundary. As the diffusion constant D is measured independently and a can be estimated experimentally by direct observation, there are in principle no adjustable parameters. It is actually impossible to get $p_2 < p < p_1$ over the whole range of τ_{off} values. This could be due to many factors including non-isotropic diffusion, complex electric-field geometry that cannot be described by a single parameter a and three-dimensional effects due to the finite thickness of the sample and the presence of the electrodes. Visual observation of the process indeed suggests that particles seldom cross the neck between electrodes during the diffusion process, the result of which is that they diffuse 'forwards'. Confirmation is provided (Fig. 4) by the good fit to the data of $p \approx 0.9 \exp(-a^2/4D\tau_{\text{off}})$, the prefactor 0.9 revealing non-isotropic diffusion.

We can learn two things from these observations. First, the drift that we observe is in semiquantitative agreement with that predicted by refs 2 and 3. The values of p shown in Fig. 4 correspond to macroscopic velocities $V = pl/(\tau_{\text{off}} + \tau_{\text{on}})$, where $l \approx 50 \mu\text{m}$ is the period of the 'Christmas tree' electrode. Plotted against τ_{off} at a fixed value of τ_{on} ($\sim 30 \text{ s}$ was in all cases sufficient to achieve trapping), V peaks at values of the order of $0.2 \mu\text{m s}^{-1}$.

Second, the use of this phenomenon in a separation (or selective pumping) device⁸ is obvious, as the velocities depend significantly on the particle size. In fact, the selectivity is much better than suggested by Fig. 4 because for each particle size the optimum frequency has been chosen, whereas in practical circumstances one would use the same frequency for all particles. Such devices are in principle well suited for viruses, DNA molecules and complete chromosomes, because their size falls within the range investigated here. The 'Christmas tree' structure was chosen for its visual convenience, but a better design of the electrodes could also improve their performance.

Our experiment could be thought of as a simple mimic of motor protein assemblies, which have recently been given much theoretical attention^{4,6,7,9,10}. It has been shown that if the motor protein evolves between at least two states in which it interacts differently with the asymmetric filament (actin or tubulin), and if the transition from one state to the other is triggered by an external agent such as adenosine triphosphate (ATP), then the motor will actively move along the filament⁴. In our experiment the electrodes provide the asymmetric

FIG. 3 Pictures of the population of colloidal particles during an on/off cycle as seen by fluorescence microscopy. Scale bar at lower left corner, 10 μm . *a*, Trapping of the 0.4- μm particles under the action of a 5 V, 500 kHz voltage. *b*, Dispersion of the particles after diffusion in the 'off' state for 24 s. *c*, Snapshot of a late stage of the retrapping process, after 18 s in the subsequent 'on' state. Most of the particles are already trapped, but in the middle of the picture are four particles on their way to the trapping site situated on the right of the picture. The sample thickness of $\sim 3 \mu\text{m}$ (as measured either directly by focusing successively on the upper and lower boundaries of the sample, or by measuring the final diameter of a squeezed polystyrene sphere of initial diameter 15 μm) was kept small enough to prevent charge injection instabilities^{1,2}.

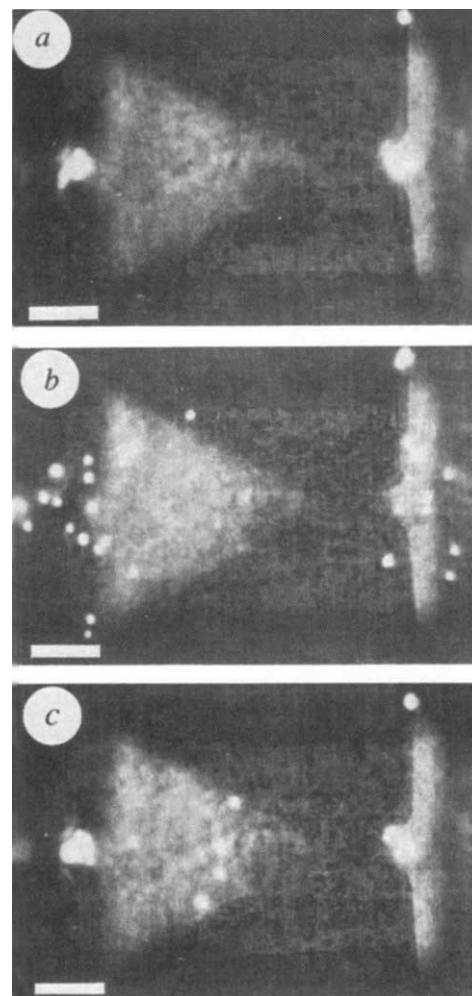
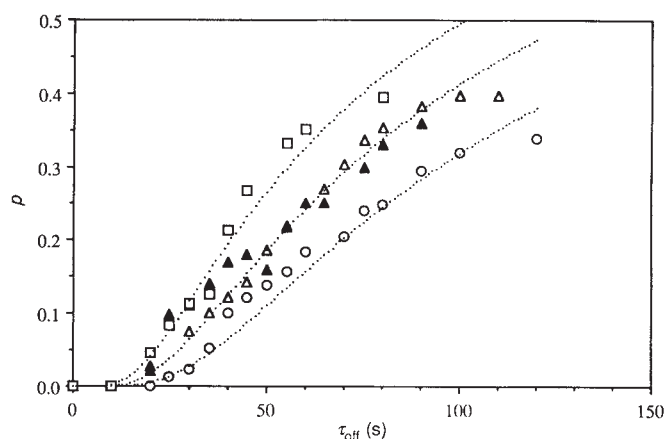


FIG. 4 Proportion p of particles diffusing from one trap to the attraction zone of the next during an 'off' (field-free) period, as a function of the time lapse τ_{off} of this period. The different plots correspond to different particle sizes and frequencies of the electric field: 0.25 μm and 800 kHz (open squares); 0.4 μm and 500 kHz (open triangles); 1.0 μm and 60 kHz (open circles). The solid triangles correspond to the direct measurement of the proportion of particles actually captured by the next trap over a cycle. The dotted lines correspond to $p \approx 0.9 \exp(-a'^2/4D\tau_{\text{off}})$ with a' equal to 19, 18 and 14 μm for particles of diameter 0.25, 0.4 and 1.0 μm respectively. The frequencies were chosen in such a way that the trapping occurred slightly after the neck on the wide side. This favoured efficient pumping.



environment, and the 'on' and 'off' periods mimic the two states. The externally imposed regular switch between 'on' and 'off' maintains the system out of equilibrium, as the ATP does in biological systems. Note that if we scale our observed velocities with a diffusion coefficient of the order of

$10^{-11} \text{ cm}^2 \text{ s}^{-1}$, even though significantly smaller than in ref. 11, and the filament period of 8 nm (in the case of tubulin) while keeping the same asymmetry, we find optimal velocities of the order of $1 \mu\text{m s}^{-1}$ which is the same order of magnitude found for dynein/tubulin systems. $\square$

Received 21 March; accepted 28 June 1994.

- Curie, P. *J. Phys. (Paris) III* **3**, 393 (1894).
- Ajdari, A. thesis, Univ. Paris 6 (1992).
- Ajdari, A. & Prost, J. *C.r. hebdo. Séanc. Acad. Sci., Paris II* **315**, 1635–1639 (1992).
- Prost, J., Chauwin, J.-F., Peliti, L. & Ajdari, A. *Phys. Rev. Lett.* **72**, 2652–2655 (1994).
- Feynman, R. P., Leighton, R. B. & Sands, M. in *The Feynman Lectures on Physics* Vol. 1 Ch. 46 (Addison-Wesley, Reading, Massachusetts, 1966).
- Astumian, R. D. & Bier, M. *Phys. Rev. Lett.* **72**, 1766–1769 (1994).

- Peskin, C. S., Ermentrout, G. B. & Oster, G. F. in *Cell Mechanics and Cellular Engineering* (ed. Mow, V. et al.) (Springer, New York, 1994).
 - Ajdari, A., Lewiner, J., Prost, J. & Viovy, J.-L. French Patent No. 9311346 (1993).
 - Magnasco, M. O. *Phys. Rev. Lett.* **71**, 1477–1481 (1993).
 - Doering, C. R., Horsthemke, W. & Riordan, J. *Phys. Rev. Lett.* **72**, 2984–2987 (1994).
 - Cordova, N. J., Vale, R. D. & Oster, G. F. in *Biologically Inspired Physics* (ed. Peliti, L.) 207–215 (Plenum, New York, 1991).
 - Felici, N. *C.r. hebdo. Séanc. Acad. Sci., Paris, Série B* **273**, 1004–1007 (1971).
- ACKNOWLEDGEMENTS. We thank P. Barois and J. L. Viovy for their help and advice.

Design and synthesis of a heterogeneous asymmetric catalyst

Kam T. Wan & Mark E. Davis

Division of Chemistry and Chemical Engineering,
California Institute of Technology, Pasadena, California 91125, USA

THE need for enantiomerically pure pharmaceuticals and insecticides drives the search for new catalysts that can perform asymmetric reactions with high enantiomeric selectivity. Many chiral transformations are accomplished using homogeneous catalysts¹. Recently we reported² a heterogeneous catalyst for asymmetric hydrogenation reactions, in which an organometallic catalytic complex is held in a film of water on a porous hydrophilic support, while the reactants and products remain within a hydrophobic organic phase. Such systems provide a large interfacial area between the catalytic phase and the solvent, and should facilitate the separation of product from catalyst. Our previous system² gave selectivities of only about 70% enantiomeric excess, however, whereas values of at least 95% are needed for practical utility. Here we describe a modified process in which an enantiomeric excess of 96% is obtained for the asymmetric reduction of 2-(6'-methoxy-2'-naphthyl)acrylic acid to the commercially important anti-inflammatory agent naproxen. We use ethylene glycol as the hydrophilic liquid phase containing the chiral catalyst, which effectively prevents the leaching of the complex into the organic phase. Our results show that these supported-phase asymmetric catalytic systems have the potential to become commercially viable.

Among the criteria necessary for a successful heterogeneous catalyst are high activity and enantioselectivity, without leaching of catalytic material into the product phase. Using these targets, we set out to design a true, heterogeneous, asymmetric catalyst. We began by using the concept of supported aqueous-phase catalysis³. This class of catalyst consists of a water-soluble organometallic complex supported in a thin film of water that resides on a high-surface-area hydrophilic solid. The reason for choosing the supported aqueous-phase catalyst as the starting point for the design is that it provides very active and selective heterogeneous analogues of known homogeneous catalysts. For example, the hydroformylation of terminal olefins proceeds with such catalysts at ~25% of the rate observed in homogeneous solution, at the same aldehyde selectivity and with no loss of the organometallic complex into the organic phase^{4,5}. We used the asymmetric reduction of 2-arylacrylic acids as the model reaction both because asymmetric reduction can be used to produce a wide variety of chiral functional groups and because 2-arylpropionic acids (for example, ibuprofen and naproxen) are high-value pharmaceutical products. It is known that ruthenium complexes of 2,2'-bis(diphenylphosphino)-1,1'-binaphthyl (BINAP) are efficient homogeneous catalysts for the asymmetric reduction of **1** to naproxen^{6,7} (Fig. 1). We developed a water-soluble analogue of BINAP by direct sulphonation⁸, and created a supported aqueous-phase catalyst comprising a ruthenium(II)-sulphonated-BINAP complex^{2,9}.

During our earlier investigations² it became apparent that the loss of a chloro ligand from the ruthenium through hydrolysis was detrimental to the enantioselectivity. The retention of the chloro ligand probably provides the molecular configuration necessary to facilitate a more enantioselective binding with the substrate. A similar loss in enantioselectivity on removal of a Ru-Cl bond was found in the Ru-BINAP system in an organic solvent¹⁰. It is known that water in a supported aqueous phase is capable of hydrolysing metal-ligand bonds¹¹. Thus the non-trivial final steps in the design process were to specify: (1) a

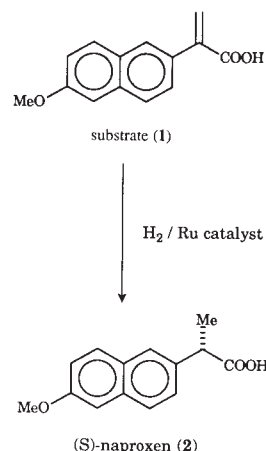


FIG. 1 Reduction of substrate (**1**) to naproxen (**2**).

non-volatile hydrophilic liquid that replaces water but that can still immobilize the water-soluble complex without cleavage of Ru-Cl bond; (2) a hydrophobic organic solvent immiscible with the components of the supported film, but miscible with reactants and products. As the highly polar ethylene glycol is not miscible with most organic solvents, it was chosen as a substitute for the aqueous phase in the original supported aqueous-phase system and used as the supported hydrophilic liquid (catalytic phase). To maintain the integrity of the ethylene glycol film during the reaction, a rather nonpolar 1:1 mixture of cyclohexane and chloroform was selected as the hydrophobic organic phase (reactant/product phase). The function of the chloroform is to enhance the substrate solubility.

The catalyst was synthesized and comprised a thin film of [Ru(BINAP-4SO₃Na)(C₆H₆)Cl]Cl (ref. 9) and ethylene glycol on a controlled-pore glass (CPG-240) that has an average pore size of 242 Å and 120/200 mesh size ((1–2.5) × 10⁻⁵ mol Ru per g catalyst). The asymmetric reduction of **1** in 5 ml of a 1:1 mixture of chloroform and cyclohexane was accomplished at a 1/3 ratio of 30–100, a hydrogen pressure of 94–101 MPa and a temperature of 276–297 K. At 276 K, the enantiomeric excess at 100% conversion of **1** is 96% for either the heterogeneous catalyst or the homogeneous system of **3** in neat methanol.

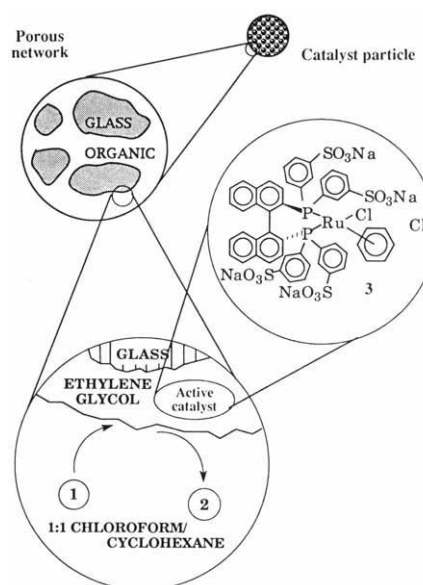


FIG. 2 Schematic diagram of the designed heterogeneous catalyst.

Additionally, the initial turnover frequencies for these catalyst systems at 297 K are 41 and 131 h⁻¹ for the heterogeneous and homogeneous catalysts, respectively. The high reaction rate and enantiomeric excess of this new heterogeneous catalyst are accomplished without loss of ruthenium to the organic phase at a detection limit of 32 p.p.b. The high enantiomeric excess obtained from this new catalyst (96% as opposed to the 70% reported previously²) makes this system of practical interest.

For long-term stability, the heterogeneous catalyst must remain assembled under the reaction conditions. To test for this type of stability, we ascertained whether the components could self-assemble into the heterogeneous catalyst configuration. The ruthenium complex **3**, dissolved in ethylene glycol, and CGP-240 were loaded separately into a reactor with **1** dissolved in 1:1 chloroform/cyclohexane. The reactor was pressurized with 97.5 Mpa H₂ and stirred (at 350 r.p.m.) at room temperature for 2 h. A control experiment was carried out in a similar manner except that no CGP-240 was added. Complete conversion of **1** was observed when CPG-240 was added, whereas <2% conversion was found in the control experiment. Additionally, the CPG-240 collected after completion of the reaction was found to contain the ruthenium complex; the bulk organic phase was colourless. These results indicate that, under the reaction conditions, the individual components of the heterogeneous catalyst self-assemble and are more stable in the heterogeneous configuration than separated. Therefore the reverse process (that is, the separation of the solution and complex from the support) is not likely under reaction conditions. These experiments also indirectly imply that the reaction takes place at the liquid-liquid interface. The limited interfacial area of the catalyst solution (in the absence of CPG-240) that is immiscible with the bulk organic phase is probably responsible for the lack of activity in the control experiment.

One of the unique features of this new heterogeneous, asymmetric catalyst, which is also one of the reasons for its success, is the method of immobilization. Rather than immobilizing the active organometallic complex through a covalent bond to a supporting surface—which usually leads to a significant loss in activity¹²—here the phase containing the organometallic complex and ethylene glycol is immobilized onto the support. At the molecular level, this method of immobilization yields a heterogeneous catalyst that is basically the same as its homogeneous analogue, thus allowing for high enantioselectivity and activity.

Although the term "catalyst design" has been used in many contexts, we believe the present example is a real demonstration of catalyst design. By successful implementation of prescribed steps we have been able to synthesize a true, heterogeneous (no leaching of Ru at a detection limit of 32 p.p.b.), asymmetric catalyst that shows high activity (one-third of the homogeneous rate) and enantioselectivity (96% enantiomeric excess). The strategies outlined here are general, and we hope that they can be applied to the development of other heterogeneous, asymmetric catalysts. □

A climate model study of indirect radiative forcing by anthropogenic sulphate aerosols

A. Jones, D. L. Roberts & A. Slingo

Hadley Centre for Climate Prediction and Research,
Meteorological Office, London Road, Bracknell,
Berkshire RG12 2SY, UK

ANTHROPOGENIC sulphate aerosols are believed to affect the radiation budget of the Earth in two ways. Through the *direct* effect they scatter solar radiation back to space, producing a radiative forcing whose global annual mean has been estimated to lie in the range -0.3 to -0.9 W m⁻² (refs 1-3). This is significant compared to the longwave forcing due to increases in anthropogenic trace gases since the beginning of the industrial era, estimated at +2 to +2.5 W m⁻² (ref. 4). Aerosols also have an *indirect* effect, altering the distribution and concentration of cloud condensation nuclei (CCN) and hence the number density and size distribution of cloud droplets, thus affecting the solar radiative characteristics of clouds^{5,6}. This is harder to quantify than the direct effect, because it depends on complex and poorly understood interactions between aerosols, CCN and cloud properties. Here we use sulphate aerosol data derived from a three-dimensional chemical transport model⁷ to estimate the indirect radiative forcing by low-level water clouds using a general circulation model. We estimate that the indirect aerosol effect at the top of the atmosphere is approximately -1.3 W m⁻² in the global annual mean. Although this value is subject to a high level of uncertainty, even if the effect is only half as large it would still exceed many estimates of the direct effect, demonstrating its potential importance in climate change.

The model used was a version of the Hadley Centre general circulation model (GCM), a configuration of the UK Meteorological Office's unified forecast/climate model⁸. The model includes a prognostic cloud scheme which predicts cloud liquid- and ice-water contents⁹, and a parametrization of the effective radius of cloud water droplets¹⁰⁻¹², which links effective radius to cloud type, aerosol concentration and cloud liquid-water content. The radiation scheme calculates the cloud radiative properties from the liquid-water content and effective radius, using parametrizations of the single-scattering properties of the cloud as input to a conventional two-stream model¹³. Cloud overlap in the vertical is random. Clouds warmer than 0 °C are assumed to consist entirely of water, and those colder than -15 °C to be entirely ice, with a smooth transition in between.

Two-dimensional distributions of total column sulphate mass loading were obtained from the model of Langner and Rodhe⁷ (the 'slow oxidation' version). Distributions of total sulphur and of sulphur from only natural sources were obtained, the difference between them being due to anthropogenic emissions. The aerosol mass distribution was calculated on the assumption that the sulphur was in the form of ammonium sulphate. To calculate distributions of aerosol particle number concentration a log-normal size distribution was assumed¹⁴, using a median particle radius of 0.05 µm and a geometric standard deviation of 2, consistent with Kiehl and Briegleb². The vertical distribution of the aerosol was approximated by assuming that half of the column-integrated mass was located in the lowest 1.5 km of the atmosphere. This yields annual average total sulphate aerosol number concentrations ranging from <100 cm⁻³ over the remote oceans to >800 cm⁻³ over central Europe (Fig. 1a). When only natural sources are considered, the maximum concentrations (of the order of 200 cm⁻³) are over the Eastern Pacific (Fig. 1b).

The relationship between the concentrations (cm⁻³) of aerosol (*A*) and the number of cloud droplets (*N*_{tot}) formed on those

Received 27 April; accepted 28 June 1994.

- Parshall, G. W. & Nugent, W. A. *CHEMTECH* March, 184-190 (1988); *CHEMTECH* May, 314-320 (1988); *CHEMTECH* June, 376-383 (1988).
- Wan, K. T. & Davis, M. E. *J. Catal.* **148**, 1-8 (1994).
- Arhancet, J. P., Davis, M. E., Merola, J. S. & Hanson, B. E. *Nature* **339**, 454-455 (1988).
- Davis, M. E. *Chemtech* August, 498-502 (1992).
- Arhancet, J. P., Davis, M. E. & Hanson, B. E. *J. Catal.* **121**, 327-339 (1990); **129**, 94-99; 100-105 (1991).
- Noyori, R. & Takaya, H. *J. org. Chem.* **52**, 3174-3176 (1987).
- Chan, A. S. C., Laneman, S. A. & Miller, R. E. in *Selectivity in Catalysis* Ch. 3 (eds Davis, M. E. & Suib, S. L.) 27-39 (Symp. Ser. 517, American Chemical Soc., Washington DC, 1993).
- Wan, K. T. & Davis, M. E. *J. chem. Soc., chem. Commun.* 1262-1264 (1993).
- Wan, K. T. & Davis, M. E. *Tetrahedron: Asymmetry* **4**, 2461-2468 (1993).
- Hoke, J. B., Hollis, L. S. & Stern, E. W. *J. organomet. Chem.* **455**, 193-196 (1993).
- Guo, I., Hanson, B. E., Toth, I. & Davis, M. E. *J. molec. Catal.* **70**, 363-368 (1991).
- Fache, E., Mercier, C., Pagnier, N., Despeyroux, B. & Panster, P. *J. molec. Catal.* **79**, 117-131 (1993).

ACKNOWLEDGEMENTS. We thank D. Forster for a generous gift of 2-(6'-methoxy-2'-naphthyl)-acrylic acid. This work was supported by the US NSF Alan T. Waterman Award (M.E.D.).

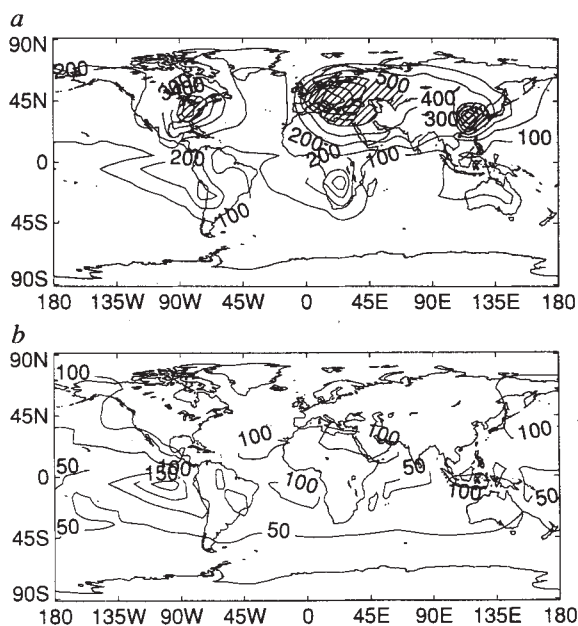


FIG. 1 *a*, Total sulphate aerosol concentration (annual mean) inferred from the Langner and Rodhe⁷ data. Global mean, 174.9 cm^{-3} . Isopleths are every 100 cm^{-3} , and values $>500 \text{ cm}^{-3}$ are hatched. *b*, As *a*, but for natural sources only, with isopleths every 50 cm^{-3} . Global mean, 75.0 cm^{-3} .

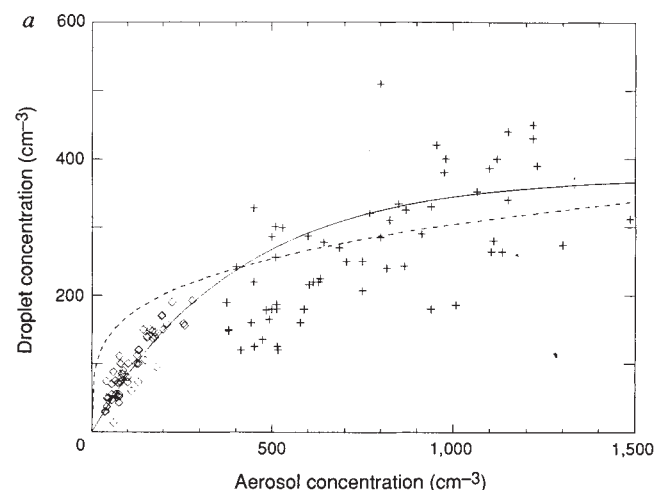
aerosols which act as CCN was based on aircraft data from a wide range of locations^{10,12}.

$$N_{\text{tot}} = 375(1 - \exp[-2.5 \times 10^{-3} A]) \quad (1)$$

This is shown as the solid line in Fig. 2*a*. Martin and Johnson^{10,12} derived separate curves for maritime and continental clouds, but equation (1) was derived to provide N_{tot} as a single continuous function of the aerosol concentration. A relation between cloud-water sulphate and N_{tot} deduced from data taken over North America (ref. 15, equation 1) was also used, assuming the same aerosol composition and size distribution, and that cloudwater and particulate sulphate concentrations are equal. This gives:

$$\log_{10}(N_{\text{tot}}) = 0.257 \log_{10}(0.122 A) + 1.95 \quad (2)$$

which is shown as the dashed line in Fig. 2*a*. At low values of



A this gives $N_{\text{tot}} > A$, which is unrealistic as the number of CCN cannot exceed the total number of aerosol. N_{tot} was therefore limited to ensure that $N_{\text{tot}} \leq A$. Due to the greater range of geographical locations underlying equation (1), this equation was preferred in the following analysis.

The cloud droplet effective radius (r_e) for stratiform and shallow ($<500 \text{ m}$ depth) convective clouds was derived using a further parametrization^{10,12}:

$$r_e = (3L/4\pi\rho\kappa N_{\text{tot}})^{1/3} \quad (3)$$

where L is the cloud liquid-water content and ρ the density of water. The value of the constant κ depends on whether the cloud is maritime ($\kappa=0.80$) or continental ($\kappa=0.67$). In this study, the assumption was made that clouds over land are continental and those elsewhere are maritime, although we recognise that continental air can be found well away from land. Figure 2*b* illustrates the behaviour of equation (3) using convenient units. For deep convective clouds, the values of r_e used were $9.5 \mu\text{m}$ for continental clouds and $13.5 \mu\text{m}$ for maritime clouds¹¹; ice cloud particles had a prescribed effective radius of $30 \mu\text{m}$. These values were the same in all integrations. For low-level layer clouds (below the $\sim 750 \text{ mb}$ level in the atmosphere), it is assumed that the droplet concentration is constant with height and that L increases monotonically from zero at cloud base to a maximum at cloud top. The droplet radius then increases with height, according to equation (3), as observed^{16,17}. The radiation scheme uses the value of r_e at cloud top, which controls the cloud radiative properties seen from space¹⁷.

Figure 3*a* shows the simulated annual-mean r_e distribution for low-level water clouds, from a control integration using the total aerosol distribution to simulate present-day conditions. Clouds with small areal coverage and very small values of r_e due to low liquid-water contents (which have little radiative significance) were excluded from this figure to avoid biasing the distribution. Mean values are $1.3 \mu\text{m}$ larger over the oceans than over land, and $2.4 \mu\text{m}$ larger in the Southern Hemisphere than in the Northern Hemisphere. These differences are in the same sense as in the satellite retrievals of Han *et al.*¹⁸, although they found a bigger ocean/land contrast ($3.3 \mu\text{m}$) and a smaller hemispheric contrast ($0.7 \mu\text{m}$). The simulation also agrees with aircraft measurements of r_e at cloud top for marine stratocumulus¹⁹, which show a roughly $3 \mu\text{m}$ increase in r_e from around the UK to the less polluted South Atlantic. Despite the many assumptions made in this study, the generally realistic nature of the r_e simulation is encouraging.

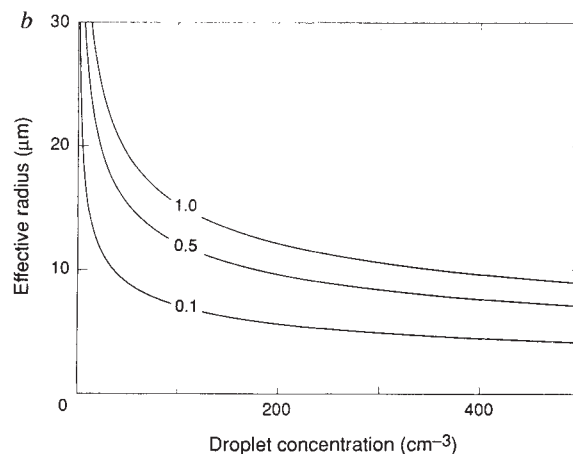


FIG. 2 *a*, Solid line, the relation (equation (1) in text) between cloud droplet and aerosol concentrations, superimposed on the observations of Martin and Johnson^{10,12}. The diamonds indicate maritime clouds, and the crosses denote continental clouds. Dashed line, equation (2), derived from Leaitch *et al.*¹⁵. *b*, The variation of cloud droplet effective radius with cloud droplet concentration according to equation (3), for continental clouds with different values of cloud liquid-water content (given on the curves in gm^{-3}).

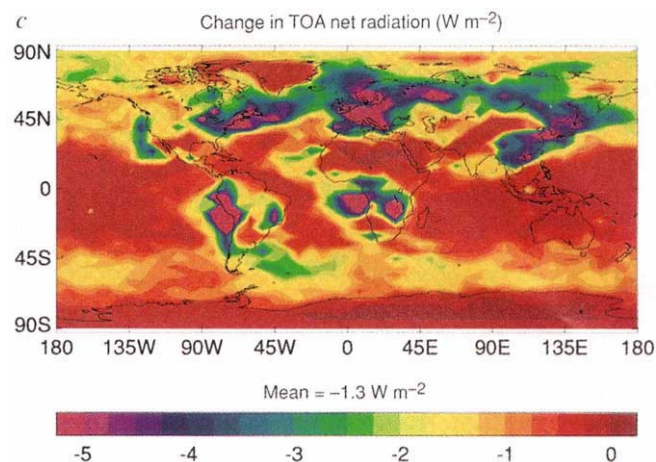
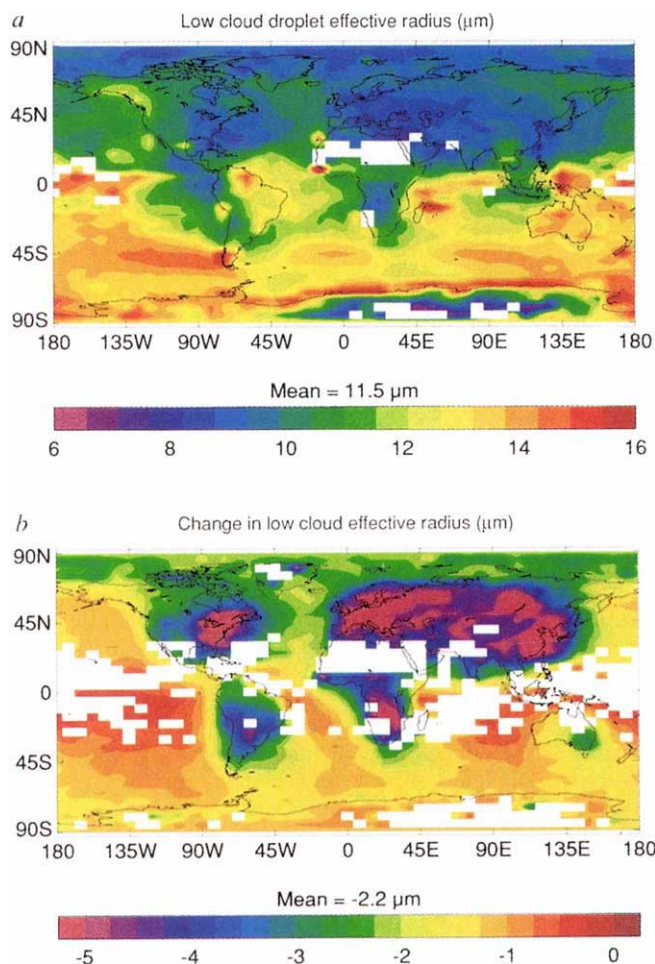


FIG. 3 a, Hadley Centre GCM simulation of annual mean distribution of low-level cloud droplet effective radius at cloud top. The blank areas indicate regions where there were no low clouds during the integration. b, Annual-mean composite of the instantaneous change in low cloud droplet effective radius due to changing from natural-only to total aerosol concentration. Blank areas indicate regions of no low cloud during any of the sampling periods. c, Distribution of the indirect radiative forcing due to the change in aerosol concentration, in terms of the change in the top of the atmosphere (TOA) net radiation.

The indirect forcing was estimated by performing a series of single-timestep calculations with the GCM, using initial conditions from the first day of each month of the control integration. For each day, the calculations were performed at 3-h intervals. The results were composited to ensure proper sampling of both the diurnal and annual cycles. The calculations were made once with the total aerosol concentration (Fig. 1a) and once with only the natural concentration (Fig. 1b). The difference isolates the indirect radiative forcing in the absence of feedbacks, which may be compared directly with the forcings from other perturbations.

The mean change in low-cloud r_e from the single-timestep calculations is shown in Fig. 3b. There is a general decrease in r_e throughout most of the Northern Hemisphere, and over most land areas, being most pronounced near the major industrial regions which produce the largest amounts of aerosol.

Figure 3c shows the annual-mean change in the top of the atmosphere (TOA) net radiation due to the indirect effect of anthropogenic sulphate aerosols. The cooling is greatest where there are large changes in the droplet effective radius and where there is sufficient low cloud which is unobscured by higher-level cloud. The calculated current anthropogenic indirect forcing of -1.26 W m^{-2} is roughly half the size of the forcing due to increases in trace gases to date; the mean Northern Hemisphere forcing is -1.54 W m^{-2} , compared with -0.97 W m^{-2} in the Southern Hemisphere. The global-mean value is comparable with some of the simple estimates of Charlson *et al.*²⁰, and is somewhat more than twice the value estimated by Kaufman and Chou using a two-dimensional climate model²¹. Recent GCM numerical experiments by Boucher and Rodhe²² gave a global-mean forcing in the range -0.65 to -1.35 W m^{-2} , although their forcing is more concentrated in the Northern Hemisphere. Our value is substantially larger than the direct radiative effect of

-0.3 W m^{-2} estimated by Kiehl and Briegleb² using the same aerosol size distribution. If the estimates of the direct and indirect effects are combined, it suggests that aerosols have a substantial overall cooling effect. This cooling is relatively localized, which has implications for the detection of anthropogenic greenhouse warming.

If the alternative relation between N_{tot} and A (based on equation (2)) is used, a value of -1.25 W m^{-2} is obtained for the indirect effect, with a similar geographical distribution. The similarity to our previous estimate is mainly due to the restriction on N_{tot} mentioned earlier. Other uncertainties could lead to a much wider spread of estimates. These uncertainties include details of the sulphur chemistry⁷, the use of the sulphate mass loadings to infer A , the use of the same size distribution for both total and natural aerosols, the generality of equations (1)–(3) and the dependence of the results on the quality of the GCM cloud simulations. The 'slow oxidation' version of Langner and Rodhe's model provides a conservative estimate of the direct aerosol effect (H. Rodhe, personal communication), but the 'fast oxidation' version would be necessary to obtain a conservative estimate of the effect of sulphur chemistry on the indirect effect. We have used the 'slow oxidation' version here to enable us to compare our results directly with Langner and Rodhe's findings, but the details of the sulphur chemistry would be unlikely to change our qualitative conclusions. Processes not addressed here include the impact of changing particle size on precipitation development and cloud evolution²³. This emphasises the need for more work on this question, including comprehensive observational studies.

Our study has excluded non-sulphate aerosols, which may significantly contribute to CCN concentrations in areas influenced by anthropogenic emissions^{24,25}. However, this should not

result in a large underestimate of the radiative forcing, because of the nonlinearity of the response to increasing CCN concentrations (assuming that the areas affected by anthropogenic sulphate and non-sulphate emissions are similar). In contrast, if natural sources of non-sulphate CCN are important then our figures will tend to overestimate the indirect radiative forcing.

The indirect forcing depends both on the aerosol loading and on the susceptibility of the clouds to changes in aerosol^{19,26}. This is highest for clouds in clean, maritime airmasses. Figure 3c thus shows significant contributions from maritime regions affected by the anthropogenic emissions. For example, over the southern oceans the clouds have a high susceptibility, which amplifies the importance of the modest anthropogenic aerosol loadings. This

effect would have been strongest at the start of the industrial era, when aerosol loadings were minimal and clouds over the whole globe were most susceptible. The indirect effect would therefore have increased quickly as emissions developed, but in the future it may saturate if the susceptibility of all clouds becomes negligible. In contrast, the direct effect increases linearly with the aerosol loading, when the optical depth is small. The indirect effect may thus have dominated early on, but if sulphur emissions continue to grow, in time the direct effect may pre-dominate. However, if sulphur emissions are reduced in the future, for example to meet concerns about acid precipitation, both the direct and the indirect effects will diminish. This could enhance the global impact of greenhouse-gas warming²⁷. □

Received 7 February; accepted 1 July 1994.

1. Charlson, R. J., Langner, J., Rodhe, H., Leovy, C. B. & Warren, S. G. *Tellus* **43**, 152–163 (1991).
2. Kiehl, J. T. & Briegleb, B. P. *Science* **260**, 311–314 (1993).
3. Taylor, K. E. & Penner, J. E. *Nature* **369**, 734–737 (1994).
4. Houghton, J. T., Jenkins, G. J. & Ephraums, J. J. (eds) *IPCC Climate Change: The IPCC Scientific Assessment* (Cambridge Univ. Press, 1990).
5. Twomey, S. A. *Atmos. Envir.* **8**, 1251–1256 (1974).
6. Twomey, S. A., Piepgrass, M. & Wolfe, T. L. *Tellus* **36B**, 356–366 (1984).
7. Langner, J. & Rodhe, H. *J. Atmos. Chem.* **13**, 225–263 (1991).
8. Cullen, M. J. P. *Met. Mag.* **122**, 81–94 (1993).
9. Smith, R. N. B. *Q. Jl. R. met. Soc.* **116**, 435–460 (1990).
10. Martin, G. M. & Johnson, D. W. in *Proc. 11th Int. Conf. Clouds and Precipitation*, Montreal 158–161 (AES, Downsview, 1992).
11. Bower, K. N. & Choulaton, T. W. *Atmos. Res.* **27**, 305–339 (1992).
12. Martin, G. M., Johnson, D. W. & Spice, A. J. *atmos. Sci.* **51**, 1823–1842 (1994).
13. Slingo, A. J. *atmos. Sci.* **46**, 1419–1427 (1989).
14. d'Almeida, G. A., Koepke, P. & Shettle, E. P. *Atmospheric Aerosols: Global Climatology and Radiative Characteristics* (Deepak, Hampton, Virginia, 1991).

15. Leaitch, W. R., Isaac, G. A., Strapp, J. W., Banic, C. M. & Wiebe, H. A. *J. geophys. Res.* **97**, 2463–2474 (1992).
16. Nicholls, S. Q. *Jl. R. Met. Soc.* **110**, 783–820 (1984).
17. Nakajima, T. & King, M. D. *J. Atmos. Sci.* **47**, 1878–1893 (1990).
18. Han, Q., Rossow, W. B. & Lacis, A. A. *J. Clim.* **7**, 465–497 (1994).
19. Taylor, J. P. & McHaffie, A. J. *atmos. Sci.* **51**, 1298–1306 (1994).
20. Charlson, R. J. et al. *Science* **255**, 423–430 (1992).
21. Kaufman, Y. J. & Chou, M.-D. *J. Clim.* **6**, 1241–1252 (1993).
22. Boucher, O. & Rodhe, H. *Department of Meteorology Report CM-83* (Stockholm Univ. 1994).
23. Albrecht, B. A. *Science* **245**, 1227–1230 (1989).
24. Novakov, T. & Penner, J. E. *Nature* **365**, 823–826 (1993).
25. Hegg, D. A., Ferek, R. J. & Hobbs, P. V. *J. geophys. Res.* **98**, 14887–14894 (1993).
26. Plattnick, S. & Twomey, S. A. *J. of appl. Meteor.* **33**, 334–347 (1994).
27. Wigley, T. M. L. *Nature* **349**, 503–506 (1991).

ACKNOWLEDGEMENTS. We thank H. Rodhe for providing the aerosol data, and our colleagues for many helpful comments. This work was supported by the UK Department of the Environment and a consortium of oil companies within the International Petroleum Industries Environmental Conservation Association (IPIECA).

Cooler estimates of Cretaceous temperatures

Bruce W. Sellwood*, Greg. D. Price* & Paul J. Valdes†

* Postgraduate Research Institute for Sedimentology, The University, Whiteknights, Reading RG6 2AB, UK

† Department of Meteorology, The University, Earley Gate, Reading RG6 2AU, UK

THE Cretaceous period is thought to have been warmer than the present^{1–3}, with higher concentrations of atmospheric greenhouse gases such as carbon dioxide⁴. It has therefore been suggested⁵ that this time period could be used by modellers as an analogue for future climate change. But the Cretaceous Equator-to-Pole temperature gradient was flatter than today's, leading some to suggest that Cretaceous climate arose from a combination of factors, with higher atmospheric carbon dioxide concentrations leading to general warming, and other factors, such as increased ocean heat transport, leading to flattening of the latitudinal temperature gradient. Here we report new records of ocean palaeotemperature for Cenomanian sites in the Atlantic and Pacific oceans which, together with a re-evaluation of published data, cast doubt on the idea that the Cretaceous period was generally warmer. These data confirm that the latitudinal temperature gradient was flatter, but suggest that the global mean temperature was much cooler than previously believed, with minimum mean equatorial temperatures close to present values and polar temperatures close to 0 °C. In the light of these findings, the climatic role of atmospheric carbon dioxide in determining Cretaceous climate is unclear, suggesting that the Cretaceous cannot be used as an analogue for future climate change.

Concepts of Cretaceous equability are based, in part, on circumstantial geological evidence such as the widespread distribu-

tions of fossil organisms interpreted to have been intolerant of cold conditions (for example, reefs which extended 5° to 15° of latitude poleward of their present range) and sedimentary facies characteristic of particular climate regimes^{6,7}. Such data have been interpreted to reflect a poleward shift of the 21 °C isotherm by at least 5° of latitude⁸, and globally averaged temperatures 6–14 °C higher than today. Cretaceous continental configuration has been estimated to produce an average global warming of 4.8 °C above present-day values, and high-latitude warming of 10 °C (ref. 9). It has been further suggested that a combination of both palaeogeographic effects and a higher atmospheric CO₂ content could produce a rise in polar temperatures in excess of 10 °C (ref. 10). Such temperature values have been extensively used in general circulation model (GCM) experiments for the Cretaceous period⁸. Palaeobotanical data show that Cretaceous plant productivity was concentrated in middle to high latitudes and rain forests, dominated by deciduous conifers, grew in seasonally dark polar coastal areas^{11,12}. Nonetheless, the likelihood of a wholly ice-free Cretaceous Earth has been seriously questioned¹³, even though direct evidence of major continental ice-sheets and permafrost areas are as yet unrecorded⁶.

Marine palaeotemperatures are determined from measurements of the oxygen isotopic composition of unaltered shells^{2,7}. Of particular importance are well-dated planktic foraminifera¹⁴, and the new interpretation presented here is partly based upon data gathered from such fossils picked from seven Ocean Drilling Project (ODP, formerly DSDP) sites (Table 1). Additional data come from other fossils (for example, belemnites) from a wide variety of localities spanning a broad range of palaeolatitudes. However, the ecology of such extinct organisms is not fully known, particularly in terms of their depth ranges and migration behaviour¹⁵, and in addition even belemnites have often been shown to have suffered some degree of diagenetic alteration^{16,17}. Our newly determined palaeotemperatures have been combined with previously published data selected as being well constrained, both in terms of the location from which samples had been collected, and their age (Albian–Cenomanian)^{16,18–22}.

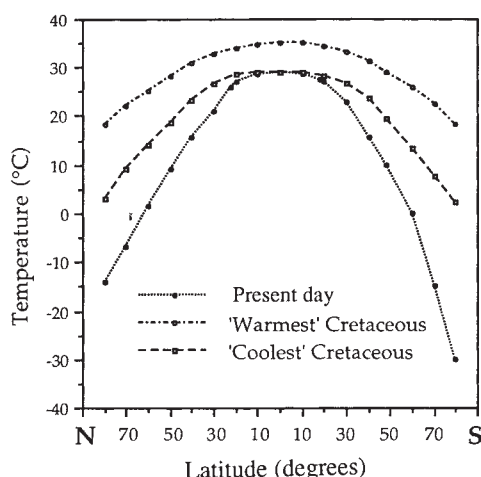


FIG. 1. Previously published estimates of latitudinal maximum and minimum temperature gradients for the mid-Cretaceous plotted for comparison against the present day mean annual surface temperature (simplified from Barron²).

All of the $\delta^{18}\text{O}$ data, including those previously published, have been calculated using the equation of Epstein *et al.*²³ and Craig²⁴, modified by Anderson and Arthur²⁵:

$$T(^{\circ}\text{C}) = 16.0 - 4.14(\delta_c - \delta_w) + 0.13(\delta_c - \delta_w)^2$$

where T is temperature, δ_c is the $\delta^{18}\text{O}$ ratio in carbonate, and δ_w is the $\delta^{18}\text{O}$ ratio of sea water. In calculating palaeotemperatures it has been assumed that Cretaceous sea water was largely free from the influence of major icecaps and a value of $\delta^{18}\text{O}$ of -1.2‰ with respect to the PDB standard (equivalent to -1‰ SMOW) has been used²⁶. This explains why our recalculated values are cooler than previous determinations of palaeotemperature using the pre-1975 data (the publication date of ref. 26).

The results presented here represent a significant advance on previously published latitudinal palaeotemperature gradients^{2,11} (Fig. 1) which have been based upon fewer locality points, significantly at low palaeolatitudes, and in which it is not always clear to what extent the effects of burial diagenesis had been considered. In addition, this study targeted a specific time-slice

within the Cretaceous period and each palaeotemperature determination is based on samples that have been evaluated rigorously in terms of possible diagenetic effects. In our study the planktic foraminifera *Rotalipora* and *Hedbergella* were selected, being easily identified and more readily assessed in terms of possible diagenetic effects. Separations involved the careful picking of each sample (a single species being separated wherever possible) from appropriate depths in the wells. *Hedbergella* is interpreted to have inhabited near-surface waters ($\sim 100\text{ m}$)²⁷, and likely to give temperatures closest to those of the surface. Foraminifera, and other shells, were investigated using standard methods of analysis and correction^{28,29}. Stable isotopic analyses were done using a VG Isogas SIRA Series II mass spectrometer, on samples averaging 2 mg in weight. Reproducibility for $\delta^{18}\text{O}$ was better than $\pm 0.1\text{‰}$ in each case. The samples were also investigated under the scanning electron microscope (SEM), and cathodoluminescence observations were also made on each analysed sample. In addition, concentrations of Mn, Sr, Mg and Fe were determined, by inductively coupled plasma spectrometry analysis on phosphoric acid residues, to check for diagenetic alteration. Only those specimens which were diagenetically unaltered were used for our palaeotemperature study (that is, non-luminescent and containing $<200\text{ p.p.m. Mn}$, $600\text{--}1,500\text{ p.p.m. Sr}$, and with neither cement nor matrix contamination visible under the SEM). Foraminiferan separates considered to consist of diagenetically altered specimens were also analysed, for comparative purposes, and were generally found to have more depleted $\delta^{18}\text{O}$ values than the unaltered ones.

Recent shallow-dwelling planktic foraminifera yield isotopically-derived temperatures $0\text{--}5^{\circ}\text{C}$ cooler than those of the surface waters, whereas temperatures cooler by $0\text{--}9^{\circ}\text{C}$ are provided by intermediate and deeper dwelling foraminifera³⁰. Our Cretaceous temperatures must be regarded as minimum estimates, as must the palaeolatitude temperature gradient curve derived both from them and from the other $\delta^{18}\text{O}$ data cited in the text (Fig. 2).

Palaeoecological factors have been invoked to explain palaeotemperatures, derived from belemnites, $4\text{--}9^{\circ}\text{C}$ lower than those determined for ammonites in the same Jurassic shales³¹. Our results show temperature values for belemnites no cooler than those derived from planktic foraminifera, suggesting that they may both be used to provide estimates of minimum surface temperatures, a view also adopted by other workers³². Despite such shortcomings, and allowing a possible increase of 5°C for these effects, a latitudinal temperature range is obtained which is sig-

TABLE 1. Data from fossilized planktic foraminifera

ODP/DSDP site	Location	Age	Foraminifera analysed	Palaeolatitude	$\delta^{18}\text{O}$ (PDB)	Temperature ($^{\circ}\text{C}$)
Site 144	Off Guiana, North Atlantic	Late Cenomanian–Early Turonian	<i>Hedbergella</i> sp.	5° N	-3.24	25.9
Site 540	Gulf of Mexico, North Atlantic	L. Albian–Mid Cenomanian	<i>Rotalipora</i> sp.	19° N	-2.65 to -1.89	21.8
Site 547A	Off Morocco, North Atlantic	Early–Late Cenomanian	<i>Rotalipora</i> sp. <i>Hedbergella</i> sp.	22° N	-3.45 to -1.89	23.9
Site 551	Gobban Spur, North Atlantic		<i>Rotalipora</i> sp.	38° N	-1.67	18.9
Site 400	Off France, North Atlantic	Albian	<i>R. ticinensis</i>	36° N	-1.57	18.4
Site 627B	Blake Plateau, North Atlantic	Early–mid Cenomanian	Mixed Planktics	25° N	-2.43 to -2.32	21.9
Site 305	Shaksky Rise, North Pacific	Late Albian–Early Cenomanian	<i>Rotalipora</i> sp. <i>Hedbergella</i> sp.	6° S	-3.55 to -2.09	24.4
Site 463	Mid-Pacific Mountains	Late Albian–Late Cenomanian	<i>Rotalipora</i> sp.	17° S	-3.01 to -2.23	22.9

Locations of ODP/DSDP sample sites, ranges of $\delta^{18}\text{O}$ values for planktic foraminiferans, together with estimates of palaeolatitude and the average of calculated palaeotemperature. The temperature range obtained was $\pm 1.5^{\circ}\text{C}$ for each site, except Site 305 where the range was $\pm 2^{\circ}\text{C}$ because here we average our $\delta^{18}\text{O}$ data with those of ref. 21. Data for Site 400 are from ref. 22.

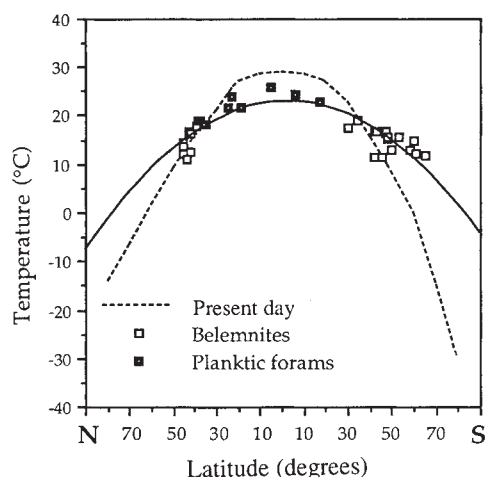


FIG. 2 Estimate of mid-Cretaceous (latest Albian-Cenomanian) latitudinal mean annual surface temperature gradient plotted against that of the present day. The Cretaceous temperature curve is a best fit based on analyses cited in Table 1 and previously published data from refs. 16, 18–22. Equatorial temperatures are significantly cooler than previous minimum estimates (for example, Fig. 1.).

effects, a latitudinal temperature range is obtained which is significantly cooler than those published hitherto, but with similar equator-to-pole gradients.

It has also been proposed, on the basis of biogenic silica accumulation, that locations such as Sites 144 (off Guiana) and 305 (Shatsky Rise) were areas of Late Cretaceous upwelling³³. If so, sea surface temperatures might have been depressed there. However, more recent work on Mesozoic oceanic silica occurrences³⁴ suggests that such upwelling interpretations are equivocal. As our samples were collected from sediments much older than those used in ref. 33, and from well below the silica-enriched horizon of the Cenomanian–Turonian oceanic anoxic event³⁵, which is recognised globally, the problems associated with upwelling probably do not apply. Palaeoceanographic reconstructions³⁶ suggest that these areas were not then environmentally restricted.

Our newly acquired Cretaceous sea surface temperatures greatly improve information on one of the boundary conditions vital in the successful running of GCM experiments. Most

models of the future predict that some warming will occur because of anthropogenic CO₂ emissions (~2 °C in the tropics given a doubling of CO₂ concentrations). But our results, taken at face value, suggest tropical temperatures during the Cretaceous period were no warmer than this, even though atmospheric CO₂ content is estimated to have been up to eight times higher than that of the present⁴. Thus increases in CO₂, and other greenhouse gases, cannot in themselves explain the very large differences that are apparent between Cretaceous and modern climate systems, and in particular the much warmer poles. Although one might speculate on the possible nature and impact of palaeocontinental distributions (and such geologically untestable phenomena as cloud cover and water vapour feedback) in generating the somewhat unexpected palaeotemperature distributions we record, we admit that our results are difficult to explain by reference to the workings of the present Earth. It is possible that estimates of atmospheric CO₂ content for the pre-Quaternary period, which are admittedly equivocal⁷, may themselves be overestimates. Nonetheless, our interpretation reinforces the probability that high-latitude upland glaciers existed, particularly over areas close to the South Pole¹³. Such a situation raises the possibility that short-term (20 kyr, 40 kyr and 100 kyr) and small-scale (<10 m) sea level changes could have been driven by orbital insolation (Milankovitch) forcing, but that the global climate effects of such forcing are likely to have been quite different when the Earth had flattened palaeotemperature gradients than when they were steep (as they are today).

Our results, suggesting Cretaceous equatorial marine temperatures comparable with those of today, also pose some intriguing questions concerning the global heat balance. During the Cretaceous period, other boundary conditions may have been very different. It is generally agreed that the solar constant is likely to have been ~1% lower than today. But the cooling effect of such a reduction could be offset by a mere doubling of the atmospheric CO₂ content above present day levels³⁷ and does little to explain the oceanic temperature distributions. Our palaeotemperature data may be partially explained if there was a higher heat transfer in Cretaceous oceans⁸. In addition, subdued continental orography could have modified atmospheric circulation patterns and at mid- to high palaeolatitudes extremes of temperature may have been ameliorated by the presence of large continental interior seaways³⁸. All these factors, and others yet to be realised, may have combined to make the Cretaceous world very different in character from that of the present. Models suggesting that future climate change might be generated by direct reference to the Cretaceous world should be viewed with extreme caution. □

Received 28 January; accepted 1 June 1994.

- Hallam, A. *J. geol. Soc. Lond.* **142**, 433–445 (1985).
- Barron, E. J. *Earth Sci. Rev.* **18**, 305–338 (1983).
- Frakes, L. A. *Climates Through Time* (Elsevier, New York, 1979).
- Berner, R. A. *Nature* **358**, 114 (1992).
- Budyko, M. I., Ronov, A. B. & Yanshin, A. L. *History of the Earth's Atmosphere* (Springer, Berlin, 1987).
- Sellwood, B. W. & Price, G. D. *Phil. Trans. R. Soc. B* **341**, 225–233 (1993).
- Crowley, T. G. & North, G. R. *Palaeoclimatology* (Oxford Univ. Press, 1991).
- Barron, E. J., Fawcett, P. J., Pollard, D. & Thompson, S. *Phil. Trans. R. Soc. B* **341**, 307–316 (1993).
- Barron, E. J. & Washington, W. M. *J. geophys. Res.* **89**, 1267–1279 (1984).
- Schneider, S. H., Thompson, S. L. & Barron, E. J. in *The Carbon Cycle and Atmospheric CO₂: Natural Variations Archaean to Present* (eds Sundquist, E. T. & Broecker, W. S.) 554–560 (American Geophysical Union, Washington DC, 1985).
- Spicer, R. A. & Corfield, R. M. *Geol. Mag.* **129**, 169–180 (1992).
- Spicer, R. A., Rees, P., McA. & Chapman, J. L. *Phil. Trans. R. Soc. B* **341**, 277–286 (1993).
- Francis, J. E. & Frakes, L. A. in *Sedimentology Review 1* (ed. Wright, V. P.) 17–30 (Blackwell, Oxford, 1993).
- Marshall, J. D. *Geol. Mag.* **129**, 143–160 (1992).
- Doyle, P. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **92**, 207–216 (1992).
- Stevens, G. R. & Clayton, R. N. *N. Z. J. Geol. Geophys.* **14**, 829–897 (1971).
- Sælen, G. *Palaeontol.* **32**, 765–798 (1989).
- Lowenstam, H. A. & Epstein, S. *J. Geol.* **62**, 207–248 (1954).
- Bowen, R. J. *Paleont.* **35**, 1077–1084 (1961).
- Ditchfield, P. W., Marshall, J. D. & Pirrie, D. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **107**, 79–101 (1994).
- Douglas, R. G. & Savin, S. M. in *Init. Rep. DSDP Leg 32*, 509–520 (1975).

- Léttle, R., Grazzini, C. V. & Pierre, C. in *Init. Rep. DSDP Leg 48*, 741–755 (1979).
- Epstein, S., Buchsbaum, R., Lowenstam, H. A. & Urey, H. C. *Geol. Soc. Am. Bull.* **64**, 1315–1326 (1953).
- Craig, H. in *Stable Isotopes in Oceanographic Studies and Palaeotemperatures* (ed. Tongiorgi, E.) 161–182 (Consiglio Nazionale delle Ricerche, Pisa, 1965).
- Anderson, T. F. & Arthur, M. A. 1.1–1.151 (Short Course No. 10, Society of Economic Paleontologists and Mineralogists, Tulsa, 1983).
- Shackleton, N. J. & Kennett, J. P. *Init. Rep. DSDP Leg 29*, 743–755 (1975).
- Leckie, R. M. *Micropaleontology* **33**, 164–176 (1987).
- McCrea, J. M. *J. chem. Phys.* **18**, 849–857 (1950).
- Craig, H. *Geochim. cosmochim. Acta* **12**, 133–149 (1957).
- Savin, S. M. & Douglas, R. G. *Geol. Soc. Am. Bull.* **84**, 2327–2342 (1973).
- Anderson, T. F., Popp, B. N., Williams, A. C., Ho, L.-Z. & Hudson, J. D. *J. geol. Soc. Lond.* **151**, 125–138 (1994).
- Barrera, E., Huber, B. T., Savin, S. M. & Webb, P. N. *Paleoceanography* **2**, 21–47 (1987).
- Miskell, K. J., Brass, G. W. & Harrison, C. G. A. *Bull. Am. Ass. Petrol. Geol.* **69**, 996–1012 (1985).
- Murray, R. W., Jones, D. L. & Buchholtz von Brink, M. R. *Geology* **20**, 271–274 (1992).
- Schlanger, S. O., Arthur, M. A., Jenkyns, H. C. & Scholle, P. A. in *Marine Petroleum Source Rocks* (eds Brooks, J. & Fleet, A. J.) 371–399 (Spec. Publ. No. 26, Geological Society of London, 1987).
- Barron, E. J. & Peterson, W. H. *Science* **244**, 684–686 (1989).
- Kasting, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **75**, 83–95 (1989).
- Valdes, P. J., Sellwood, B. W. & Price, G. D. *Palaeoclimatology* (in the press).

ACKNOWLEDGEMENTS. We are grateful for the help we received from ODP staff at ECR (Lamont-Doherty) and WCR (Scripps Institution for Oceanography) during our sampling of cores; we thank M. Isaacs for the stable isotope analyses. This work was supported by the UK NERC as part of the Mesozoic climate modelling project.

Cloning and functional expression of a rat heart K_{ATP} channel

M. L. J. Ashford*, C. T. Bond†, T. A. Blair†
& J. P. Adelman†‡

* Department of Pharmacology, University of Cambridge, Tennis Court Road, Cambridge CB2 1QJ, UK

† Vollum Institute, Oregon Health Sciences University, Portland, Oregon 97201, USA

POTASSIUM channels that are ATP-sensitive (K_{ATP}) couple membrane potential to the metabolic status of the cell. K_{ATP} channels are inhibited by intracellular ATP and are stimulated by intracellular nucleotide diphosphates¹. K_{ATP} channels are important regulators of secretory processes and muscle contraction, and are targets for therapeutic treatment of type II diabetes by the inhibitory sulphonylureas² and for hypertension by activators such as pinacidil³. In cardiac tissue, K_{ATP} channels are central regulators of post-ischaemic cardioprotection^{4,5}. Electrophysiological and pharmacological characteristics vary among K_{ATP} channels recorded from diverse tissues suggesting extensive molecular heterogeneity¹. A complementary DNA encoding a K_{ATP} channel was isolated from rat heart using the polymerase chain reaction. We report here that the expressed channels possess all of the essential features of native cardiac K_{ATP} channels, including sensitivity to intracellular nucleotides. In addition the cloned channels are activated by the potassium channel opener, pinacidil, but are not inhibited by the sulphonylurea, glibenclamide.

Two overlapping sets of degenerate oligonucleotides directed to the consensus pore sequence of cloned potassium channels were used in the polymerase chain reaction (PCR), using cDNA synthesized from rat heart messenger RNA as substrate. When the reaction products were subcloned and the nucleotide and predicted amino-acid sequences examined, one clone appeared to encode a novel potassium channel. On the basis of this sequence a full-length coding sequence, rc K_{ATP} -1 was obtained.

The 417 amino-acid sequence of rc K_{ATP} -1 (Fig. 1a) displays the hallmarks of the inward rectifier family⁶⁻⁹. The N and C termini are predicted to be within the cell connected by two hydrophobic, probably transmembrane segments which flank a region with strong homology to other potassium channel pore sequences.

Northern blot analysis of rc K_{ATP} -1 mRNA (Fig. 1b) detected a positively hybridizing mRNA of ~3.8 kilobases (kb) in heart ventricle and atrium; this band was not detected at high stringency in several insulin-secreting islet cell lines. In other experiments using reduced hybridization stringency, an additional band was detected in islet cell mRNA which represents an islet cell specific K_{ATP} sequence (unpublished observations). A broader tissue distribution study was done using the PCR, showing that rc K_{ATP} -1 mRNA is expressed in a variety of peripheral and central nervous system tissues (Fig. 1c).

The rc K_{ATP} -1 was heterologously expressed by transient transfection in either a human (HEK293) or hamster (BHK21) kidney epithelial cell line. Inside-out membrane patches excised from cells transfected with rc K_{ATP} -1 contained a distinctive conductance not present in mock-transfected cells, or cells transfected with the same vector expressing β -galactosidase. In symmetrical (140 mM) K^+ , these channels showed bursting patterns of activity interspersed with long silent periods (Fig. 2a). The single-channel currents exhibited some inward rectification even in the absence of internal Mg^{2+} . However, the degree of rectification was enhanced by increasing the concentration of

FIG. 1 a, Predicted amino-acid sequence for rc K_{ATP} -1 and the human cardiac K_{ATP} , and comparison to ROMK1, IRK1 and GIRK1. Identical residues are boxed. Dots represent gaps introduced to optimize sequence alignments. rc K_{ATP} -1 shows 73% similarity (identical and conservative substitutions) to GIRK, 70% similarity to IRK, and 64% similarity to ROMK. Transmembrane segments and the pore region are overlined; asterisks denote a potential glycosylation site. b, Northern analysis of rc K_{ATP} -1 mRNA. Poly (A)⁺ RNAs from rat heart ventricle (5 μ g), atrium (0.5 μ g), and a rat insulinoma cell line, RinM5F (5 μ g). A ~3.8 kb band is detected primarily in ventricle, but is also expressed in atrium (note differences in amounts of mRNA loaded). At high stringency, no bands were detected in RinM5F mRNA. c, Tissue distribution of rc K_{ATP} -1 mRNA by PCR.

METHODS. a, Poly (A)⁺ RNA isolated from rat heart (Fast Track RNA isolation kit, Invitrogen) was reverse transcribed (Superscript, BRL) using a bipartite synthetic primer. Amplification was achieved by nested PCR using degenerate oligonucleotide pools directed to the consensus amino-acid sequence of the pore region of cloned potassium channels as described in ref. 22. 5' and 3' RACE PCR²³ were done to obtain the full coding sequence. The nucleotide and predicted amino-acid sequence of rc K_{ATP} -1 have been deposited with Genbank. b, For Northern blot analysis, poly (A)⁺ RNAs were size fractionated on a 1% agarose-formaldehyde gel, transferred to a nylon membrane and probed with a cRNA probe containing 960 bases of rc K_{ATP} -1 coding sequence. Integrity of the RNA samples was verified by ethidium staining before transfer. c, Total RNA from the indicated tissue was reverse transcribed using random primers. Amplification used unique oligonucleotides directed against 5' untranslated and N-terminal coding sequences. Reaction products were prepared as a Southern blot and probed with a specific radiolabelled oligonucleotide directed to a sequence between the amplification primers.

internal Mg^{2+} (Fig. 2b). There was no clear dependence of open-state probability (P_o) on voltage, although in two patches the open-state probability increased by about twofold with depolarization (-50 mV to $+50$ mV). The unitary slope conductance for inward currents under symmetrical 140 mM KCl conditions was 70.2 ± 2.0 pS ($n = 13$) and was independent of the Mg^{2+} concentration at the cytoplasmic surface. The currents reversed at 0 mV under symmetrical KCl conditions. In experiments with 60 mM K^+ and 80 mM Na^+ in the electrode, the potential at which single-channel currents reversed was shifted to -22.5 mV ($n = 2$) a value close to that predicted (-21.3 mV) from the Nernst equation for the potassium equilibrium potential.

Application of 1 mM Mg-ATP to the intracellular side of inside-out membrane patches containing single rc K_{ATP} -1 channels completely inhibited channel activity; this effect was readily reversed on washout (Fig. 3a), although in most experiments channel activity did not return to pre-ATP levels because of channel rundown (see below). Concentrations of ATP greater than 0.2 mM caused complete abolition of channel activity ($n = 21$) and 0.1 mM ATP reduced activity to $25.7 \pm 10.1\%$ of control ($n = 3$), indicating that the inhibitory half-maximal concentration for ATP is less than 100 μ M, in agreement with values reported in the literature for native channels^{10,11}. The effects of ATP are not due to phosphorylation because the nonhydrolysable ATP analogue AMP-PNP was an equally effective and reversible inhibitor ($n = 4$; Fig. 3b). Application of Mg^{2+} -free ATP (0.2 mM and 1.0 mM; $n = 2$) or Mg-ADP (0.5 mM; $n = 2$) also induced channel closure in a reversible manner (not shown).

Native K_{ATP} channel activity spontaneously declines after excision of the membrane patch. This characteristic, termed rundown, shows a variable time course in cardiac cells ranging from just a few seconds to several minutes¹². Heterologously expressed rc K_{ATP} -1 channels also undergo rundown over the same time periods. Like native K_{ATP} channels, the cloned channels are reactivated by application of nucleotide diphosphates to the intracellular side of the patch. When rc K_{ATP} -1 channels were allowed to rundown following patch excision, application of 10 mM UDP evoked some channel activity (Fig. 3c). In the absence of Mg^{2+} , UDP did not reactivate rc K_{ATP} -1 channels,

‡ To whom correspondence should be addressed.

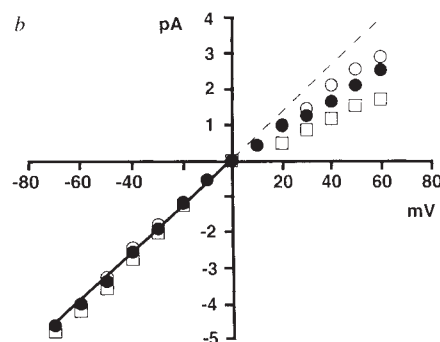
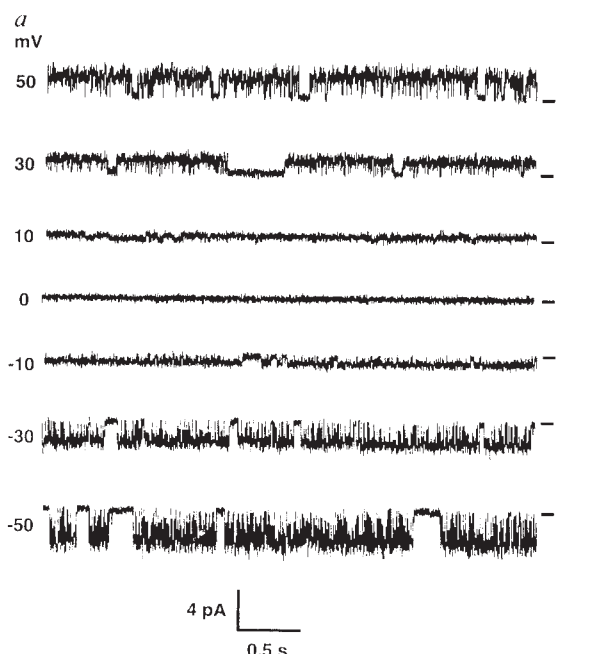
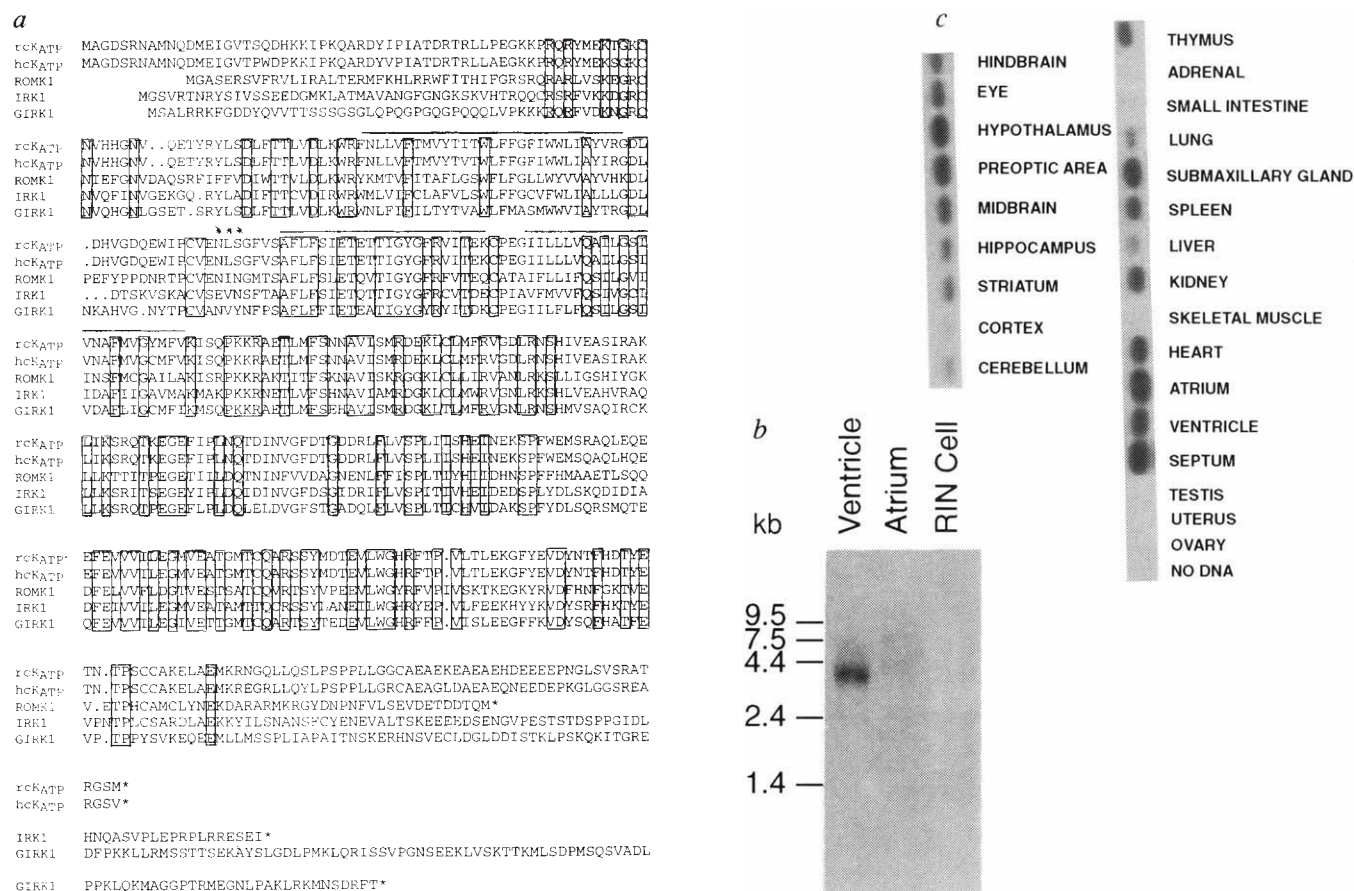


FIG. 2 Single-channel recordings from inside-out membrane patches excised from HEK293 cells transfected with $prcK_{ATP-1}$. **a**, Current tracings from a patch containing a single active channel recorded in symmetrical 140 mM K^+ solutions with 0.3 mM free Mg^{2+} at the cytoplasmic side at various membrane potentials (V_m). The closed state of the channel at each membrane potential is indicated by the bar. For display the currents were low-pass filtered at a cutoff frequency of 500 Hz. **b**, Unitary current amplitudes derived from amplitude histograms plotted against membrane potential for currents obtained in the absence (empty circles) and the presence of 0.3 mM (filled circles) and

2.0 mM (empty squares) $MgCl_2$. The straight line shows the best-fit to the 0.3 mM Mg^{2+} data and has a slope conductance of 68.1 pS. The deviation of the outward current amplitude from the ohmic conductance (broken line) is dependent on the concentration of Mg^{2+} . Note that in the absence of Mg^{2+} the relation still deviates from ohmic conduction and this may reflect intrinsic rectification of the channel or inability to remove all the Mg^{2+} near the internal surface of the excised patch.

METHODS. The coding sequence of rcK_{ATP-1} was subcloned into the tissue culture expression vector pcDNAneo (Invitrogen) which uses the CMV (cytomegalovirus) promoter to direct expression. This plasmid, $prcK_{ATP-1}$ was introduced by lipofection (Lipofectin and Opti-MEM, BRL) according to manufacturer's instructions, into either freshly thawed HEK293 or BHK21 cells. Following lipofection, cells were washed and fed with standard growth media (DMEM supplemented with 10% FCS, Gibco-BRL), and incubated at 37 °C under 5% CO_2 for 24 h. Single-channel currents were recorded from excised inside-out membrane patches²⁴ using an Axopatch 1B amplifier at 22–25 °C. The pipette solution was 140 mM KCl, 1 mM $MgCl_2$, 1 mM $CaCl_2$, 10 mM HEPES (pH 7.4). The bath solution was 140 mM KCl, 0.3 mM $MgCl_2$, 2 mM $CaCl_2$, 5 mM EGTA, 10 mM HEPES (pH 7.2). Details of single-channel recording, data collection, and analysis have been described previously^{21,25}.

whereas in the presence of 2 mM Mg^{2+} , UDP induced some brief channel bursts, although reactivation did not restore full channel activity ($n=5$). However, the rescue of rundown rcK_{ATP-1} channels was much more robust and long-lasting if 100 μM ATP was coapplied with Mg-UDP (Fig. 3d). The combination of ATP and nucleotide diphosphates resulted in very long periods of intense channel activity even following extensive washing to remove the nucleotides ($n=9$). The mean value of P_0 following rundown was 0.003 ± 0.001 ($n=9$) and following addition of Mg-UDP/ATP, P_0 increased to 0.400 ± 0.048 ($n=9$) although in some patches activation was much more pronounced as the number of functional channels also increased (Fig. 3d). After washout following rescue with Mg-UDP/ATP,

rcK_{ATP-1} channels were less sensitive to inhibition by ATP, with 1 mM causing only incomplete inhibition (Fig. 3d).

Like native K_{ATP} channels, the single channel kinetics of the cloned channels consist of long bursts of short duration channel openings and closings separated by longer closed periods (Figs 2a and 3a-c). Open time and closed time distributions of Mg-UDP/ATP-induced rcK_{ATP-1} channel activity were constructed from single-channel records and each fitted with a single exponential, resulting in time constants for open and closed times of 2.05 ± 0.24 ms and 0.31 ± 0.03 ms ($n=7$), respectively. These data are in good agreement with data for native cardiac K_{ATP} channels^{13,14}.

Cardiac K_{ATP} channels are inhibited by sulphonylureas, the

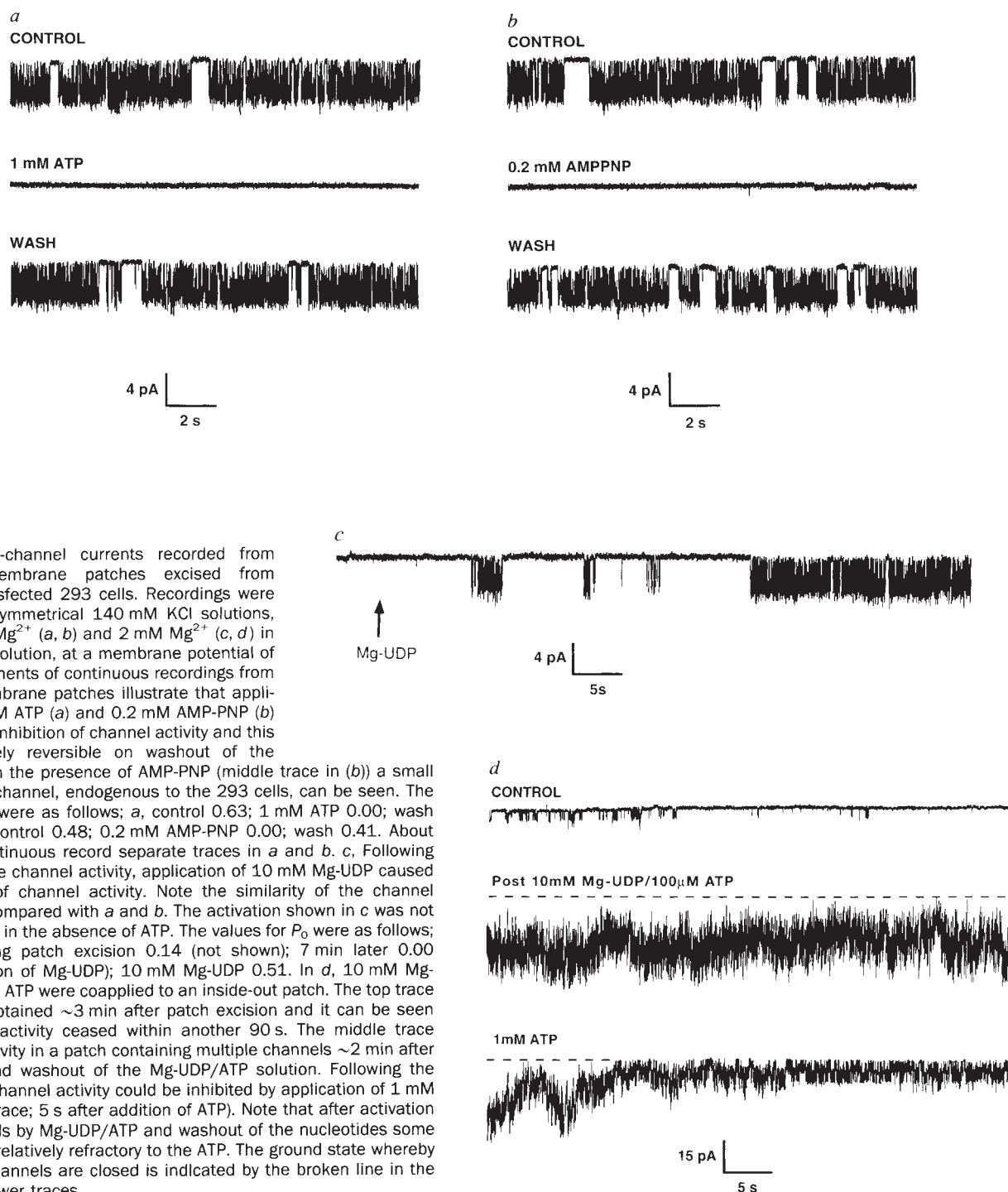


FIG. 3 Single-channel currents recorded from inside-out membrane patches excised from $prcK_{ATP-1}$ -transfected 293 cells. Recordings were made using symmetrical 140 mM KCl solutions, with 0.3 mM Mg^{2+} (a, b) and 2 mM Mg^{2+} (c, d) in the cytosolic solution, at a membrane potential of -60 mV. Segments of continuous recordings from separate membrane patches illustrate that application of 1 mM ATP (a) and 0.2 mM AMP-PNP (b) induced total inhibition of channel activity and this was completely reversible on washout of the nucleotides. In the presence of AMP-PNP (middle trace in (b)) a small conductance channel, endogenous to the 293 cells, can be seen. The values for P_0 were as follows; a, control 0.63; 1 mM ATP 0.00; wash 0.65 and b, control 0.48; 0.2 mM AMP-PNP 0.00; wash 0.41. About 5–10 s of continuous record separate traces in a and b. c, Following rundown of the channel activity, application of 10 mM Mg-UDP caused re-activation of channel activity. Note the similarity of the channel records in c compared with a and b. The activation shown in c was not well sustained in the absence of ATP. The values for P_0 were as follows; 2 min following patch excision 0.14 (not shown); 7 min later 0.00 (before addition of Mg-UDP); 10 mM Mg-UDP 0.51. In d, 10 mM Mg-UDP + 100 μM ATP were coapplied to an inside-out patch. The top trace shows data obtained ~ 3 min after patch excision and it can be seen that channel activity ceased within another 90 s. The middle trace shows the activity in a patch containing multiple channels ~ 2 min after application and washout of the Mg-UDP/ATP solution. Following the reactivation, channel activity could be inhibited by application of 1 mM ATP (bottom trace; 5 s after addition of ATP). Note that after activation of the channels by Mg-UDP/ATP and washout of the nucleotides some channels are relatively refractory to the ATP. The ground state whereby all rcK_{ATP-1} channels are closed is indicated by the broken line in the middle and lower traces.

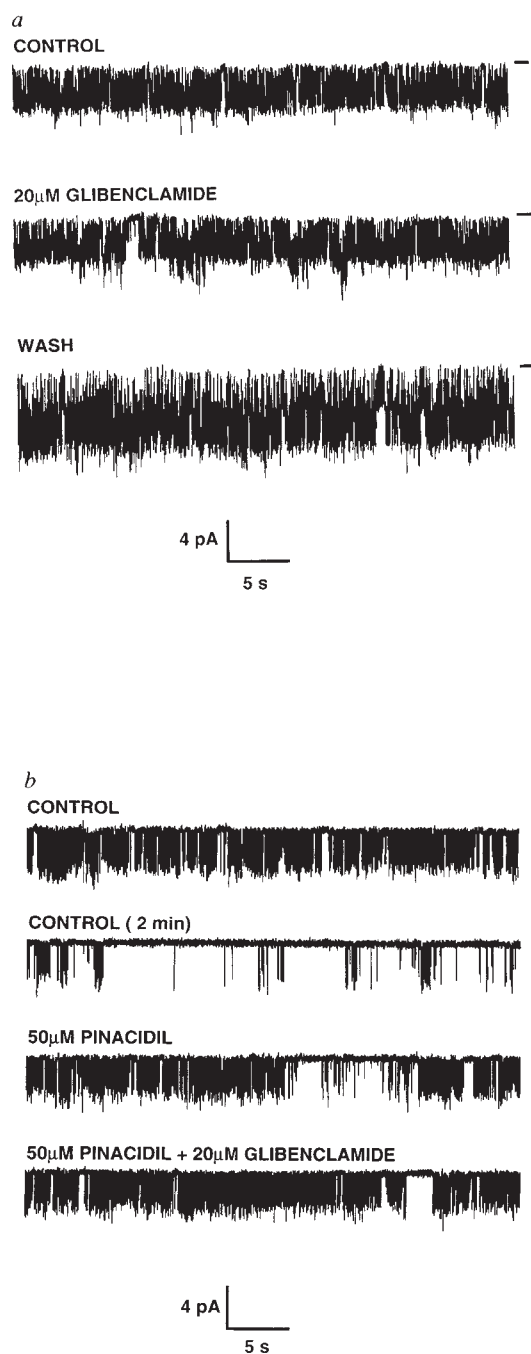


FIG. 4 $\text{rcK}_{\text{ATP}}-1$ channels expressed in either 293 or BHK cells were not sensitive to inhibition by the sulphonylurea, glibenclamide, but were activated by the potassium channel opener, pinacidil. *a*, $\text{rcK}_{\text{ATP}}-1$ channel activity recorded following activation by Mg-UDP/ATP at a membrane potential of -60 mV in the absence (top trace), presence of $20 \mu\text{M}$ glibenclamide (middle trace), and following washout of the drug (lower trace). Note that following washout of glibenclamide a second $\text{rcK}_{\text{ATP}}-1$ channel became active in the patch. The closed state of the channel is indicated by the bar. The values for P_0 were as follows; control 0.69; $20 \mu\text{M}$ glibenclamide 0.70; wash 0.57. *b*, Recordings from a patch held at a membrane potential of -60 mV. Following excision from the cell, channel activity was initially high (upper trace) and was allowed to rundown (second trace; 2 min after first trace). Application of $50 \mu\text{M}$ pinacidil in the presence of $100 \mu\text{M}$ ATP activated $\text{rcK}_{\text{ATP}}-1$ channels (third trace; 7 min later) and addition of $20 \mu\text{M}$ glibenclamide in the presence of pinacidil and ATP had no effect on channel activity (bottom trace; 90 s after third trace). The values for P_0 were as follows; control 0.38; control (2 min) 0.08; $50 \mu\text{M}$ pinacidil + $100 \mu\text{M}$ ATP 0.42; $50 \mu\text{M}$ pinacidil + $100 \mu\text{M}$ ATP + $20 \mu\text{M}$ glibenclamide 0.45.

most effective of which is glibenclamide¹⁵. Application of either $10 \mu\text{M}$ ($n=5$) or $20 \mu\text{M}$ ($n=3$) glibenclamide had no effect on the open-state probability of cloned K_{ATP} channels (Fig. 4a); for example, the normalized P_0 was 1.02 ± 0.11 ($n=4$) following addition of the $10 \mu\text{M}$ glibenclamide. Structurally diverse molecules such as pinacidil, levromakalim and nicorandil increase K_{ATP} channel activity in cardiac tissue¹⁶, an effect antagonized by glibenclamide¹⁷. We tested the action of pinacidil, a cyanoguanidine which stimulates K_{ATP} channels in the presence of intracellular ATP¹⁷. Application of pinacidil at a concentration of $30 \mu\text{M}$ ($n=3$) or $50 \mu\text{M}$ ($n=4$) in the presence of 0.1 or 0.2 mM ATP to inside-out patches containing $\text{rcK}_{\text{ATP}}-1$ channels caused channel activation; this is particularly clear following rundown (Fig. 4b). P_0 was 0.002 ± 0.002 ($n=5$) following rundown and after addition of pinacidil increased to 0.376 ± 0.082 ($n=5$). In addition, following activation of $\text{rcK}_{\text{ATP}}-1$ channels by pinacidil and ATP, concomitant application of $20 \mu\text{M}$ glibenclamide had no effect on channel activity (Fig. 4b).

The $\text{rcK}_{\text{ATP}}-1$ cDNA is a member of the inward rectifier K^+ channel family and encodes an ATP-sensitive K^+ channel with remarkable similarity in single-channel conductance, rectification, kinetics and nucleotide sensitivity, to native cardiac K_{ATP} channels. Although several nucleotides modulate $\text{rcK}_{\text{ATP}}-1$ channels, there are no obvious structural motifs which resemble consensus sequences for nucleotide binding folds¹⁸. This contrasts with ROMK1^{7,19} in which a single putative ATP-binding site was identified⁸. However, ROMK1 channel activity is not inhibited by ATP and in 20% of isolated patches high concentrations (2.5–5.0 mM) of Mg-ATP were shown to activate channel activity. Unlike native K_{ATP} channels, the cloned channels are not inhibited by micromolar concentrations of glibenclamide, although they are activated by the potassium channel opener, pinacidil. This result strongly suggests that the channel and the sulphonylurea receptor are separate molecular entities which is consistent with some recent reports^{20,21}. Thus, it appears that the sulphonylurea receptor is not a necessary constituent of K_{ATP} channels, but may modulate channel function. Coexpression of cloned K_{ATP} channels and sulphonylurea receptors will permit an understanding of the relationship between these two molecules. □

Received 3 March; accepted 24 June 1994.

1. Ashcroft, F. M. A. *Rev. Neurosci.* **11**, 97–118 (1988).
2. Gopalakrishnan, M., Janis, R. A. & Triggle, D. J. *Drug Dev. Res.* **28**, 95–127 (1993).
3. Friedel, H. A. & Brogden, R. N. *Drugs* **39**, 929–967 (1990).
4. Lynch, J. J. Jr, Sanguinetti, M. C., Kimura, S. & Bassett, A. L. *FASEB J.* **6**, 2952–2960 (1992).
5. Escande, D. & Caverio, I. *Trends pharmacol. Sci.* **13**, 269–272 (1992).
6. Kubo, Y., Baldwin, T. J., Jan, Y. N. & Jan, L. Y. *Nature* **362**, 127–133 (1993).
7. Ho, K. et al. *Nature* **362**, 31–38 (1993).
8. Dascai, N. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6596–6600 (1993).
9. Kubo, Y., Reuveny, E., Slesinger, P. A., Jan, Y. N. & Jan, L. Y. *Nature* **364**, 802–806 (1993).
10. Lederer, W. J. & Nichols, C. G. *J. Physiol.* **419**, 193–211 (1989).
11. Tung, R. T. & Kurachi, Y. *J. Physiol.* **437**, 239–256 (1991).
12. Findlay, I. *Biochim. biophys. Acta* **943**, 297–304 (1988).
13. Trube, G. & Hescheler, J. *Pflügers Arch.* **401**, 178–184 (1984).
14. Zilberter, Y., Burnashev, N., Papin, A., Portnov, V. & Khodorov, B. *Pflügers Arch.* **411**, 584–589 (1988).
15. Findlay, I. *J. Pharmacol. exp. Ther.* **261**, 540–545 (1992).
16. Edwards, G. & Weston, A. H. A. *Rev. Pharmacol. Toxicol.* **33**, 597–637 (1993).
17. Fan, Z., Nakayama, K. & Hiraoka, M. *J. Physiol.* **430**, 273–295 (1990).
18. Saraste, M., Sibbald, P. R. & Wittinghofer, A. *Trends biochem. Sci.* **15**, 430–434 (1990).
19. Aldrich, R. *Nature* **362**, 107–108 (1993).
20. Khan, R. N., Hales, C. N., Ozanne, S. E., Adogu, A. A. & Ashford, M. L. J. *Proc. R. Soc.* **253**, 225–231 (1993).
21. Lee, K., Ozanne, S. E., Hales, C. N. & Ashford, M. L. J. *Br. J. Pharmacol.* **111**, 632–640 (1994).
22. Bond, C. T. et al. *Recept. Channels* (in the press).
23. Frohman, M. A., Dush, M. K. & Martin, G. R. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8998–9002 (1988).
24. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. *Pflügers Arch.* **391**, 85–100 (1981).
25. Dempster, J. *Computer Analysis of Electrophysiological Signals* (Academic, London, 1993).

ACKNOWLEDGEMENTS. We thank our colleagues N. Marrion, R. Goodman, M. Kavanaugh and R. A. North of the Vollum Institute for enthusiasm and critical comments. J.P.A. thanks C. Miller for insightful suggestions and kibbitzing. M.L.J.A. acknowledges the support of the Parke-Davis Exchange Fellowship.

Activation of transforming growth factor- β is inhibited in transgenic apolipoprotein(a) mice

David J. Grainger, Paul R. Kemp,
Alexander C. Liu*, Richard M. Lawn*
& James C. Metcalfe

Department of Biochemistry, University of Cambridge,
Tennis Court Road, Cambridge CB2 1QW, UK

* Division of Cardiovascular Medicine,
Stanford University School of Medicine, Stanford,
California 94305, USA

A HIGH concentration of serum lipoprotein(a) is a risk factor for atherosclerosis¹⁻³. Lipoprotein(a) consists of low-density lipoprotein with the additional protein component, apolipoprotein(a), a homologue of plasminogen⁴. Lipoprotein(a) and apolipoprotein(a) enhance proliferation of human vascular smooth muscle cells (hVSMCs) in culture by inhibiting activation of plasminogen to plasmin, thus blocking the proteolytic activation of transforming growth factor- β (TGF- β)⁵, an autocrine inhibitor of hVSMC proliferation^{5,6}. The hypothesis that this pathway is a key step in atherogenesis⁵ is tested on transgenic mice expressing the human apolipoprotein(a) gene. We show here that the activation of TGF- β is inhibited in the aortic wall and serum of mice expressing apolipoprotein(a), as a consequence of apolipoprotein(a) inhibition

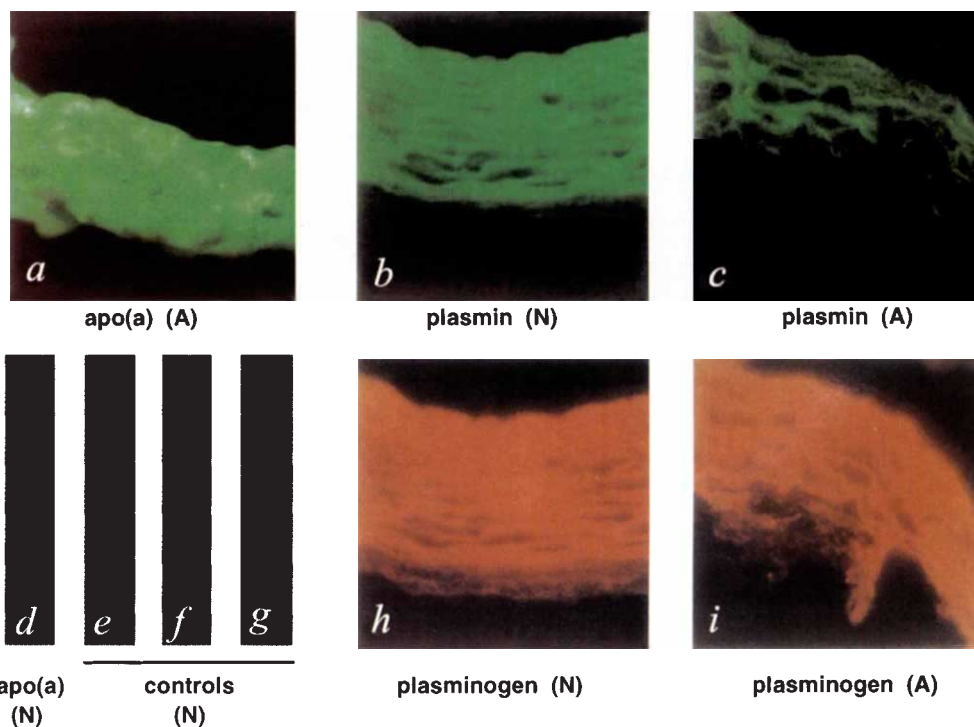
of plasminogen activation. These effects are closely correlated with VSMC activation.

Human apolipoprotein(a) (apo(a)) has been expressed in transgenic mice, a species normally lacking apo(a)^{7,8}. On a lipid-rich diet, apo(a) mice develop vascular lesions similar to fatty streak lesions in early human atherosclerosis⁷. The serum concentration of apo(a) in apo(a) mice was 3.8 ± 1.2 mg dl⁻¹ (s.e.m.; $n=14$). Immunofluorescence labelling of apo(a) in aortic sections from apo(a) mice shows foci of apo(a) accumulation near the luminal surface of the aortic wall, but apo(a) is also detected at lower concentrations throughout the media of the vessel (Fig. 1a). No apo(a) is detected in aortic sections from normal littermate mice (Fig. 1d). Plasma-derived lipoprotein(a) (Lp(a)) is known to penetrate human arteries^{9,10} and the arteries of mice injected with radiolabelled Lp(a)¹¹ or radiolabelled recombinant apo(a) (D.J.G., unpublished observations).

The inhibition of plasminogen activation by apo(a) may contribute to the development of vascular lesions by inhibiting fibrinolysis¹². Recent studies of human VSMCs in culture demonstrated that by inhibiting plasminogen activation, apo(a) also promotes smooth muscle cell proliferation and migration by suppressing the activation of TGF- β ^{5,13}. Plasmin was assayed in mouse aortic sections using the specific covalent inhibitor, α 2-antiplasmin (α 2-AP), which does not bind to plasminogen, apo(a) or other proteins in the sections (Fig. 1e, f). α 2-AP labelled with fluorescein isothiocyanate (α 2-AP-FITC) was incubated with frozen aortic sections from apo(a) mice and normal littermates. The α 2-AP-FITC was apparently aligned with the elastic laminae of normal and apo(a) mice vessels (Fig. 1b, c). There was about threefold less active plasmin in the vessel wall of the apo(a) mice than normal mice, whether the mice had been fed a lipid-rich or normal diet (Table 1).

FIG. 1 Apo(a), plasminogen and plasmin in the aortic wall. a, Apo(a) protein labelled in the aortic section of an apo(a) (A) mouse. d, No apo(a) was detected in aortic sections of normal (N) mice. b, c, Plasmin in aortic sections from a normal or apo(a) mouse labelled with α 2-AP-FITC. e, f, Controls for nonspecific fluorescence labelling: e, aortic sections from a normal mouse incubated with α 2-AP-FITC with an excess (1 mL) of plasmin, or f, with 100 μ g ml⁻¹ plasmin inhibitor, aprotinin, for 1 h before α 2-AP-FITC. h, i, Same sections as in b and c treated with tissue plasminogen activator (tPA) to activate plasminogen then labelled with α 2-AP-TRITC (note exposures were 10-fold longer for h and i than for b and c). g, tPA omitted before labelling section from a normal mouse with α 2-AP-TRITC after labelling with α 2-AP-FITC.

METHODS. Aortae from apo(a) mice and normal littermates^{7,8} were immediately frozen in liquid nitrogen, embedded, and 6- μ m frozen sections were fixed in ice-cold acetone for 90 s. All fluorescent labelling procedures were at 4 °C. For apo(a) immunolabelling, sections were incubated with TBS, 3% BSA for 30 min, then with sheep anti-human Lp(a) antibody⁷ diluted 1:1,000 in TBS, 3% BSA followed by rabbit anti-sheep IgG-FITC diluted 1:80 in TBS, 3% BSA. α 2-AP was labelled with FITC or TRITC as described²¹. Sections were incubated for 16 h with α 2-AP-FITC (1 μ g ml⁻¹), washed in TBS containing 0.2% Nonidet-P40 and 300 mM NaCl (wash buffer), incubated with 1 mg ml⁻¹ recombinant tPA in TBS for 3 h, washed and incubated with α 2-AP-TRITC (1 μ g ml⁻¹) for 16 h before washing thoroughly in wash buffer



then TBS. All fluorescence images ($\lambda_{exc}=440$ nm, $\lambda_{em}=510$ nm for FITC; and $\lambda_{exc}=490$ nm, $\lambda_{em}=580$ nm for TRITC) were obtained using $\times 400$ magnification and exposures of 5 s for apo(a), 10 s for α 2-AP-FITC and 1 s for α 2-AP-TRITC. There were no significant differences in the 2–5% autofluorescence backgrounds in sections from different mice. Images shown are representative of aortic sections from at least 8 mice in each group, aged 24 weeks, and were indistinguishable whether the mice were fed a normal or lipid-rich diet.

TABLE 1 Quantitative fluorescence data

	Normal mice		apo(a) mice	
	Normal diet	Lipid-rich	Normal diet	Lipid-rich
Plasminogen				
Total	702 ± 47	748 ± 95	789 ± 121	688 ± 133
% Plasmin	6.3 ± 1.3	6.1 ± 0.6	1.7 ± 0.7*	1.9 ± 1.2*
TGF- β				
Total	112 ± 7	95 ± 12	115 ± 1	109 ± 6
% Active	90 ± 6	90 ± 5	36 ± 3*	46 ± 8*
Osteopontin				
Intensity	1.4 ± 0.8	0.4 ± 0.1	32.3 ± 4.4*	12.6 ± 2.1*
Area	0.7 ± 0.9	1.2 ± 1.6	80.3 ± <0.1*	103 ± 31.7*

Fluorescence intensities of consecutive aortic sections from apo(a) and normal mice assayed for plasmin, plasminogen, total and active TGF- β and osteopontin. The mice were fed a normal or a lipid-rich diet from age 12 weeks until killed at 24 weeks. Four apo(a) and four normal mice were fed each diet and four sections from the aorta of each mouse were labelled (see Fig. legends). Fluorescence was quantified using an extended linear-range TV camera (Photonic Science) attached to a Nikon Diaphot inverted fluorescence microscope and Magiscan image analysis software (Joyce-Loebl). The gain control on the photomultiplier was set so that the average autofluorescence of the unlabelled vessel wall was between 2 and 5% of the full-scale intensity. For each section, fluorescence images were captured for four fields of aortic wall selected randomly under phase contrast ($\times 400$). For TGF- β and plasminogen/plasmin the average fluorescence intensity for the 16 fields of view for each aorta was calculated. The mean values for each group of mice (s.e.m.; $n=4$) are shown. For osteopontin, the vessel media was only partly labelled and only pixels with intensity values $>5\%$ of maximal intensity was included in the averaged fluorescence intensity. The number of pixels ($\times 10^{-2}$) above the threshold is shown as the area labelled for osteopontin.

* $P < 0.05$ for apo(a) mice compared with normal mice (Student's unpaired t-test).

Active plasmin in the vessel wall of apo(a) mice was decreased because apo(a) inhibited plasminogen activation rather than reduced the amount of plasminogen. To assay plasminogen, active plasmin was first labelled with $\alpha 2$ -AP-FITC as above, then the same sections were treated with 1 mg ml $^{-1}$ tissue plasminogen activator (tPA) and relabelled for plasmin using $\alpha 2$ -AP coupled to trimethylrhodamine isothiocyanate ($\alpha 2$ -AP-TRITC; Fig. 1*h, i*). Where the tPA was omitted, no further staining for active plasmin with $\alpha 2$ -AP-TRITC was observed (Fig. 1*g*). There was no significant difference in the total amounts of plasminogen in the sections from the apo(a) and normal mice, whereas about 6% of the plasminogen was activated to plasmin in the normal mice, compared with only 2% in the apo(a) mice (Table 1).

To determine whether the low plasmin concentration in the aortic wall of the apo(a) mice resulted in reduced activation of TGF- β , immunofluorescent labels were used to quantify active TGF- β and total TGF- β (active + latent). Total TGF- β in the vessel wall was assayed using an anti-TGF- β antibody which detects active and latent TGF- $\beta 1$ and TGF- $\beta 3$ with equal sensitivity and a trimethylrhodamine-labelled second antibody. TGF- β was present throughout the aortic media, apparently aligned with the elastic laminae (Fig. 2*a, b*) in sections from normal and apo(a) mice. There was no significant difference in the amount of total TGF- β present in aortic sections from normal and apo(a) mice (Table 1).

Active TGF- β was assayed in the same aortic sections as total TGF- β using a truncated extracellular domain of the type II TGF- β receptor fused to glutathione-S-transferase (R2X) and labelled with fluorescein isothiocyanate (R2X-FITC; Fig. 2*d, e*). R2X binds active, but not latent, TGF- $\beta 1$ and TGF- $\beta 3$ with equal affinity¹⁴. Control experiments demonstrated the specificity of the active and total TGF- β fluorescent labels (Fig. 2*c, f*). To determine the proportion of active TGF- β in the sections, the

ratio of fluorescent intensities of the active and total TGF- β labels was calibrated using mixtures of active and latent TGF- β (Fig. 2*i*). False-colour images of the proportion of TGF- β in the aortic wall that was active showed that the TGF- β in the aortic wall of apo(a) mice (Fig. 2*h*) is significantly less active than the TGF- β in the aortic wall of normal mice (Fig. 2*g*), whether the mice had been fed a lipid-rich or normal diet (Table 1). Furthermore, activation of TGF- β is most strongly inhibited at the sites of highest apo(a) accumulation. The dark blue (false colour) areas near the luminal surfaces of aortic sections from apo(a) mice on lipid-rich or normal diets (Fig. 3*b, c*) correspond to sites of focal apo(a) accumulation, assayed in adjacent sec-

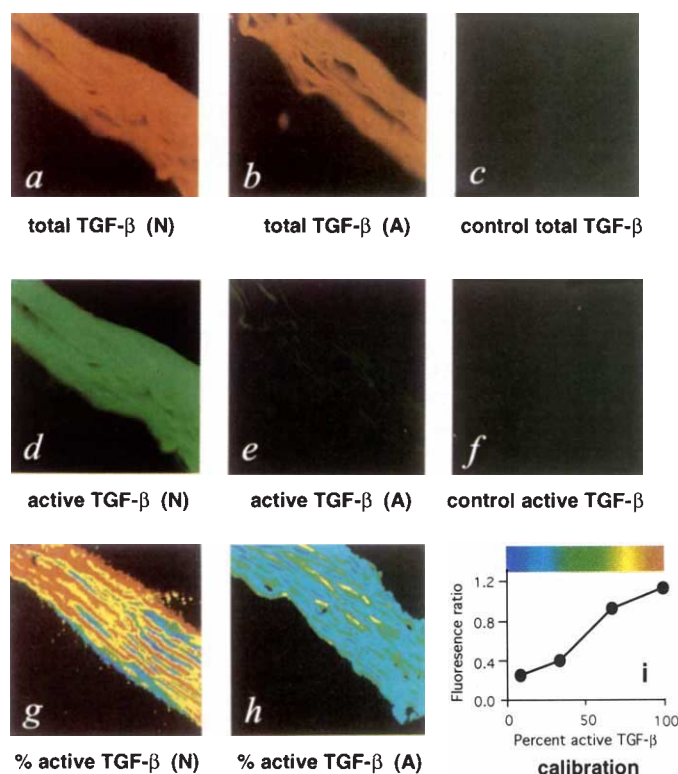
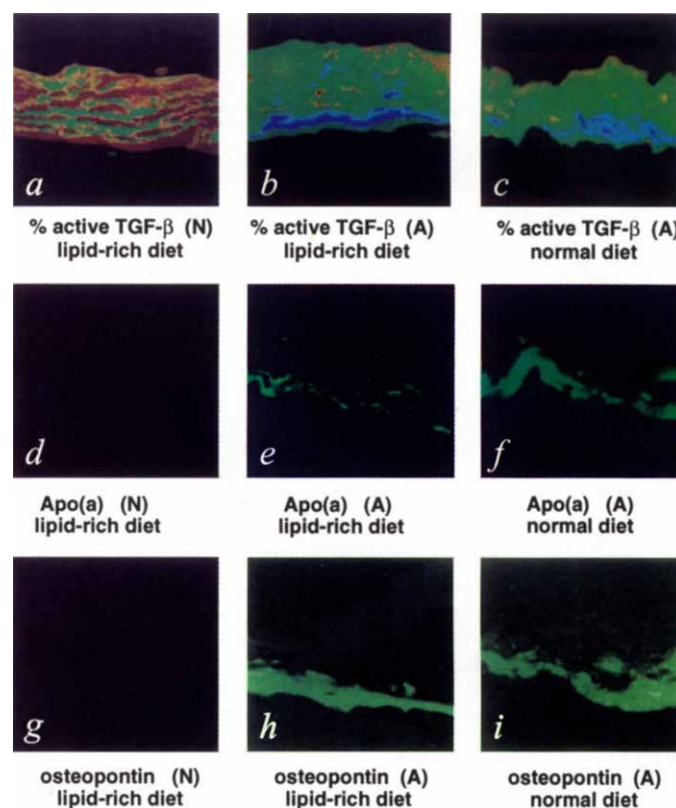


FIG. 2 Active and total TGF- β in the aortic wall. *a, b*, Total TGF- β in aortic sections from a normal (N) mouse and an apo(a) (A) mouse. *d, e*, Active TGF- β in the same aortic sections shown in *a* and *b*. *c, f*, Controls for nonspecific labelling of sections by total and active TGF- β labels: *c*, aortic section from a normal mouse incubated with the rhodamine-labelled second antibody in the absence of the primary antibody and *f*, similar section incubated with R2X-FITC in the presence of 100 ng ml $^{-1}$ TGF- $\beta 1$. *g, h*, False-colour images of the proportion of active TGF- β in the section from a normal mouse (computed from *a* and *d*) and in the section from an apo(a) mouse (computed from *b* and *e*). *i*, Calibration of the fluorescence intensity ratio of fluorescein and rhodamine labels of active and total TGF- β .

METHODS. Sections were labelled for total TGF- β for 2 h with 25 μ g ml $^{-1}$ BDA47 in TBS, 3% BSA (R&D systems), washed 3 times in TBS and incubated with goat anti-rabbit IgG conjugated with TRITC (Sigma) diluted 1:50 in TBS, 3% BSA. The same section was then washed 3 times in TBS and labelled for active TGF- β with R2X-FITC (D.J.G. *et al.*, manuscript in preparation). Sections were incubated with R2X-FITC (5 μ g ml $^{-1}$) for 16 h, then washed 3 $\times$ in PBS. To calibrate the fluorescence intensities of the active and total TGF- β labels, solutions containing various proportions of active TGF- β (6 ng ml $^{-1}$ total TGF- β) were spotted on to gelatine-polylysine-coated slides and dried. The spots were labelled for total and active TGF- β as for the aortic sections and the fluorescence intensity ratios (TRITC/FITC) determined. Fluorescent images were obtained as in Fig. 1 (5-s exposures). The images shown are typical of sections from at least 8 mice in each group and were indistinguishable whether the mice were fed a normal or lipid-rich diet.

FIG. 3 TGF- β and VSMC activation. *a–c*, False-colour images of TGF- β activation in aortic sections of mice aged 24 weeks. *a*, A normal (N) mouse fed a lipid-rich diet from age 12 weeks to 24 weeks; *b*, an apo(a) (A) mouse fed a lipid-rich diet from age 12 weeks to 24 weeks; and *c*, an apo(a) mouse fed a normal diet from birth to 24 weeks. *d–f*, Apo(a) labelling in adjacent sections to those in *a*, *b* and *c*, respectively, done as for Fig. 1*a* except that the exposure time was reduced to 1 s to show only the areas of high apo(a) concentration. *g–i*, Osteopontin immunofluorescent labelling in adjacent sections to those in *a*, *b* and *c*, respectively.

METHODS. Osteopontin was labelled with the monoclonal antibody MP11B10₁ (National Institute of Health Developmental Studies Hybridoma Bank) at 10 $\mu\text{g ml}^{-1}$ in TBS, 3% BSA for 16 h, washed 3 times in TBS and detected by incubating with goat anti-mouse IgG-FITC (Sigma F-2012; 1:50 dilution in TBS, 3% BSA) for 2 h. The images are typical of all mice examined, with at least 4 in each subgroup.



tions (Fig. 3*e,f*). These observations strongly suggest that changes in the vessel wall that are a consequence of reduced TGF- β activity will occur preferentially at the sites of focal apo(a) accumulation.

Inhibition of TGF- β activation by apo(a) was also detected in mouse serum by enzyme-linked immunosorbent assays (ELISAs) for total and active TGF- β (D.J.G. *et al.*, manuscript in preparation). In the serum of apo(a) mice the total TGF- β concentration was $14.4 \pm 4.7 \text{ ng ml}^{-1}$ ($n=21$) whereas in the normal mice it was $14.2 \pm 3.5 \text{ ng ml}^{-1}$ ($n=21$). However, the proportion of total TGF- β in the active form in the serum of apo(a) mice was $34 \pm 19\%$ compared with $92 \pm 12\%$ active TGF- β for normal mice ($P<0.005$).

To examine the effects of apo(a) on VSMCs in the apo(a) mice, aortic sections were labelled for osteopontin, a marker for activated smooth muscle cells^{15–18}. Immunofluorescent labelling of osteopontin was detected in aortic sections from apo(a) mice on a lipid-rich or normal diet (Fig. 3*h,i*; Table 1). Although a small increase in labelling for osteopontin was detected throughout the media of the aortae from apo(a) mice, very high concentrations of osteopontin labelling were colocalized with regions of focal apo(a) accumulation (Fig. 3*e,f*) and very low TGF- β activation (Fig. 3*b,c*). However, treating apo(a) mice with bromodeoxyuridine for 24 h before killing showed no significant mitotic activity in the aortic media, consistent with studies which showed that in the absence of physical injury, replication rates in atheromatous plaques are low¹⁹, reflecting the slow growth of the lesions. Areas of aortic sections from normal mice with high proportions of active TGF- β (Fig. 3*a*) did not show detectable labelling for osteopontin (Fig. 3*g*). The total areas and intensities of osteopontin labelling in the normal mouse sections were also very low compared with the apo(a) mouse sections (Table 1). The activation of the VSMCs may be a response to low local concentrations of active TGF- β . The data indicate that the presence of apo(a) leads to osteopontin expression in VSMCs in the aortic wall, whether the mice are fed a lipid-rich or normal diet. The accumulation of lipid into

the vessel wall on a lipid-rich diet might therefore be a consequence, rather than the cause, of VSMC activation. In this regard it is of interest that activated VSMCs in culture accumulate about 20-fold more lipid than contractile VSMCs²⁰.

Taken together, the data for the apo(a) mouse are consistent with the proposed pathway linking apo(a) to the inhibition of plasminogen and latent TGF- β activation. We suggest that the effects on smooth muscle cells of inhibiting TGF- β activation are responsible for the subsequent development of fatty lesions when the mice are subjected to a lipid-rich diet. □

Received 23 February; accepted 30 June 1994.

1. Utermann, G. *Science* **246**, 904–910 (1989).
2. Genest, J. *et al. Circulation* **85**, 2025–2033 (1992).
3. Schreiner, P. *et al. Arterioscl. Thromb.* **13**, 826–833 (1993).
4. McLean, J. W. *et al. Nature* **330**, 132–137 (1987).
5. Grainger, D. J. *et al. Science* **260**, 1655–1659 (1993).
6. Kirschenlohr, H. L., Metcalfe, J. C., Weissberg, P. L. & Grainger, D. J. *Am. J. Physiol.* **265**, C571–C576 (1993).
7. Lawn, R. M., Wade, D. P., Hammer, R. E., Verstuyft, J. G. & Rubin, E. M. *Nature* **360**, 670–672 (1992).
8. Chiesa, G. *et al. J. Biol. Chem.* **267**, 24369–24374 (1992).
9. Rath, M. *et al. Arteriosclerosis* **9**, 579–592 (1989).
10. Cushing, G. L. *et al. Arteriosclerosis* **9**, 593–603 (1989).
11. Kreutzer, J. *et al. Chem. Phys. Lipids* **67/68**, 175–190 (1994).
12. Miles, L. A. & Plow, E. F. *Thromb. Haemostasis* **63**, 331–335 (1990).
13. Kojima, S., Harpel, P. C. & Rifkin, D. B. *J. Cell Biol.* **113**, 1439–1445 (1991).
14. Massague, J. A. *Rev. Cell Biol.* **6**, 597–641 (1990).
15. Shanahan, C. M., Weissberg, P. L. & Metcalfe, J. C. *Circ. Res.* **73**, 193–204 (1993).
16. Giacelli, C., Bae, N., Lombardi, D., Majesky, M. & Schwartz, S. *Biochem. biophys. Res. Commun.* **177**, 867–873 (1991).
17. Giacelli, C. M. *et al. J. Clin. Invest.* **92**, 1686–1696 (1993).
18. Gadeau, A. P., Campan, M., Millet, D., Candresse, T. & Desgranges, C. *Arterioscl. Thromb.* **13**, 120–125 (1993).
19. O'Brien, E. R. *et al. Circ. Res.* **73**, 223–231 (1993).
20. Campbell, J. H., Reardon, M. F., Campbell, G. R. & Nestel, P. J. *Arteriosclerosis* **5**, 318–324 (1985).
21. Rinderknecht, H. *Experientia* **16**, 430–436 (1960).

ACKNOWLEDGEMENTS. We thank H. Lin (Whitehead Institute, Cambridge, MA) for providing a plasmid encoding the full-length TGF- β type II receptor and E. Rubin and J. Verstuyft (Lawrence Berkeley Laboratory) for sharing their expertise with the apo(a) mouse. This work was supported by a Wellcome Trust project grant to J.C.M. and a NIH program grant to R.M.L. A.C.L. holds a NRSA award from the National Institutes of Health. D.J.G. is the recipient of a Wellcome Prize Fellowship.

Major expansion of CD8⁺ T cells with a predominant V β usage during the primary immune response to HIV

Giuseppe Pantaleo*, James F. Demarest*, Hugo Soudeyns†‡, Cecilia Graziosi*, François Denis†, Joseph W. Adelsberger§, Persephone Borrow||, Michael S. Saag||, George M. Shaw#, Rafick P. Sekaly†‡** & Anthony S. Fauci*

* Laboratory of Immunoregulation, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

† Laboratory of Immunology, Institut de Recherches Cliniques de Montréal, Montréal, PQ H2W 1R7, Canada

§ Program Resources Incorporated/Dyncorp, Frederick, Maryland 21702, USA

¶ Infectious Diseases and # Hematology/Oncology, Department of Medicine, University of Alabama at Birmingham, Birmingham, Alabama 35294, USA

|| Department of Neuropharmacology, Division of Virology, The Scripps Research Institute, La Jolla, California 92037, USA

** Département de Microbiologie et d'Immunologie, Université de Montréal, Montréal, Québec H3C 3J7, Canada

‡ Department of Microbiology and Immunology, McGill University, Montreal Quebec H3A 2B4, Canada

A SIGNIFICANT proportion (up to 70%) of individuals experience an acute clinical syndrome of varying severity associated with primary infection with the human immunodeficiency virus (HIV)¹⁻⁴. We report here studies on six individuals who showed an acute HIV syndrome which generally resolved within four weeks, concomitant with a dramatic downregulation of viraemia²⁻⁵. To characterize the T-cell-mediated primary immune response to HIV, we used combined semiquantitative polymerase chain reaction assay and cytofluorometry to analyse the T-cell antigen receptor repertoire in sequential peripheral blood mononuclear cells from the patients. We found major oligoclonal expansions in a restricted set of variable-domain β -chain (V β) families. Cells expressing the expanded V β s predominantly expressed the CD8 T-cell differentiation antigen and mediated HIV-specific cytotoxicity. Major oligoclonal expansions of these CD8⁺ T lymphocytes may represent an important component of the primary immune response to viral infections and may help to clarify both the immunopathogenic and the protective mechanisms of HIV infection.

The V β repertoire in patient 1 was analysed on days 16, 20, 34 and 136 after the onset of symptoms (Fig. 1). At day 16 an unusually large proportion (40%) of peripheral blood mononuclear cells (PBMC) belonged to the V β 19 subset (Fig. 1). This proportion of V β 19⁺ cells declined over time (24.6% at day 20, and 10.2% at day 34), and by day 136 had dropped to 4.2% (Fig. 1). Similar results were obtained by cytofluorometric analysis (Fig. 2a). Cells from patient 2 showed transient increases only in V β 14 (14.8%) at day 39 and V β 21 (9%) at day 17 (Fig. 1), while those from patient 3 already expressed high levels of V β 4 by day 6 (18%) which decreased progressively (14.5%, 6.4% and 3.2% at days 14, 21 and 65, respectively); in contrast, V β 8, which was poorly expressed at day 6 (2.3%), was significantly expanded at day 65 (11.8%) (Fig. 1). Cells from patient 4 showed only a moderate, late increase in V β 22 (Fig. 1) while, in contrast, those from patient 5 showed a large increase (23%) in V β 9 at day 17 followed by a dramatic decrease (4%) by day 217 (Fig. 1); a moderate expansion in V β 1 and V β 22 occurred at early time points (Fig. 1). Finally, in cells from patient 6, V β 7 increased over time (from 5.2% at day 20 to 18% at day 35),

V β 21 increased transiently (from 3.3% at day 20 to 12.6% at day 28), and V β 16 and V β 22 increased at late time points (Fig. 1). These results clearly show that a major expansion in certain V β families occurs during primary HIV infection.

Two-colour cytofluorometric analysis, carried out to determine whether CD4⁺ or CD8⁺ T cells were involved in this V β expansion, showed that most (95%) cells expressing V β 19 were T lymphocytes expressing CD8 (Fig. 2b and c). At day 16, 48.6% CD8⁺ T cells co-expressed V β 19, whereas only 2.5% CD4⁺ T cells were V β 19⁺ (Fig. 2b and c). CD8⁺V β 19⁺ cells decreased progressively (Fig. 2b), whereas CD4⁺V β 19⁺ cells did not change significantly over time (Fig. 2c). Furthermore, three-colour cytofluorometric analysis showed that 72% of CD8⁺ T cells and 87% of CD8⁺V β 19⁺ cells co-expressed the activation marker HLA-DR (Fig. 2d, compared with only 8% of all CD4⁺ cells (Fig. 2e). These results strongly suggest that activated CD8⁺ T cells expressing the expanded V β subset are involved in the primary immune response to HIV.

The nature of the V β expansion (that is, whether oligoclonal or polyclonal) was determined by nucleotide sequencing of the V-D-J (variable-diversity-joining) segments of T-cell receptor β -chain recombinants. We determined nucleotide sequences of recombinant clones of V β 19 derived from PBMC of patient 1 collected on day 20, when CD8⁺V β 19⁺ cells were dramatically expanded (40%), and on day 136, when they were 5%. Although all the rearrangements were unique with respect to junctional diversity, there was a heavy bias in the use of the J β 2.1 segment (10 of 11 clones) (Fig. 3a). Also, the overall length of the V-D-J junction in these clones was limited to four to six amino acids, suggesting structural constraints on the V β sequence (Fig. 3a). This idea is further supported by the presence of a glutamic acid-glutamine (EQ) in single-letter amino-acid nomenclature) motif 5' of the J β region in 10 of 11 clones sequenced (Fig. 3a). At day 136, there was still a heavy bias in the use of J β 2.1 (10 of 11 clones) (Fig. 3b). More importantly, 9 of 11 recombinant clones now showed identical V-D-J rearrangements; these rearrangements were identical to that of clone 1.69 at day 20 (Fig. 3a), indicating selection over time. Our results show the oligoclonal nature of the V β 19 expansion. Similar molecular analysis of V β 4 on cells from patient 3 confirmed the oligoclonal nature of these expansions. At day 6 (V β 4 18%) there was a heavy bias in the use of J β 2.7 in 8 of 9 clones, with 6 showing identical rearrangements (data not shown). By day 65, although V β 4 had decreased to 3.2% (Fig. 1), 4 of 5 clones still showed a bias in the use of J β 2.7. Interestingly J β 2.1 and J β 2.7 (that is, the two J β segments expressed in the expanded V β 19 and V β 4 in patients 1 and 3, respectively) are the only 2 of 13 known germline J β segments which contain an EQ motif⁶. These residues 5' of the J β region are known to be critical in interactions of T-cell receptors with complexes between peptide and the major histocompatibility complex (MHC)⁷.

To determine whether CD8⁺ T lymphocytes were also involved in the moderate expansions of additional V β s observed in four of six patients (patients 2, 3, 5 and 6) (Fig. 1), three-colour cytofluorometry was carried out on PBMC from patient 3 on days 21 and 65. By day 21, 2.4% of cells were V β 8⁺ and showed equal numbers of CD4⁺ and CD8⁺ cell subsets (1.4% CD4⁺V β 8⁺ cells and 1.0% CD8⁺V β 8⁺ cells) (Fig. 3c). By day 65, V β 8⁺ cells had increased to 7.6%. This increase was due almost exclusively to an expansion in CD8⁺V β 8⁺ cells (6.3%); the percentage of CD4⁺V β 8⁺ cells was unchanged (1.3%; Fig. 3c). Most CD8⁺V β 8⁺ cells (86%) expressed HLA-DR at day 65 (Fig. 3c). These results strongly suggested that this expansion of activated CD8⁺V β 8⁺ cells was oligoclonal, and this was confirmed when nucleotide sequencing of recombinant V β 8 clones showed that at day 6 there was a highly diverse V β repertoire, with 6 of the 13 known J β families being represented (Fig. 3d), whereas at day 65 only J β 2.3, J β 2.1 and J β 1.6 were present (Fig. 3e). Notably, all the J β 2.3 clones (5 of 9 clones) had identical rearrangements (Fig. 3e). The oligoclonal nature of these expan-

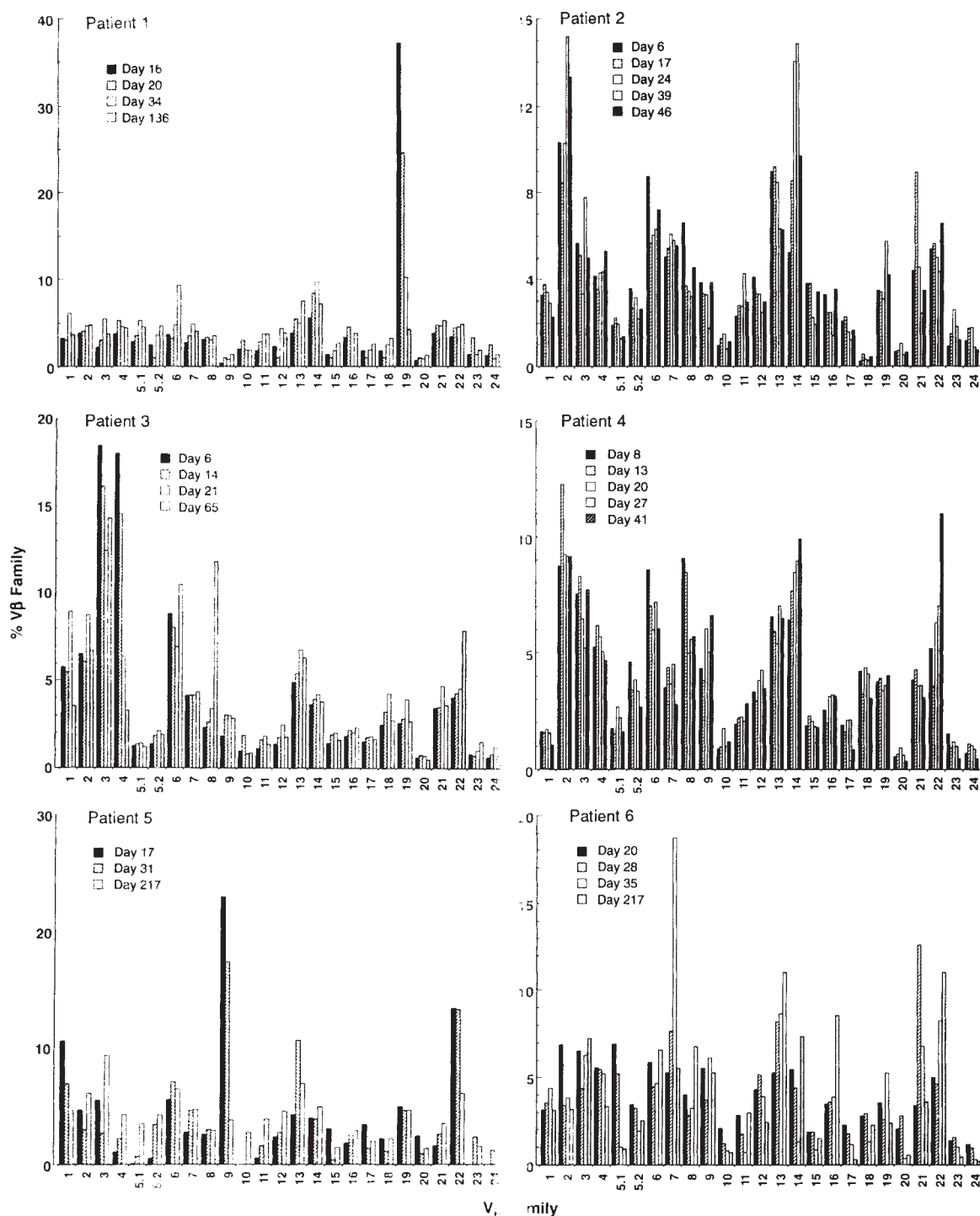


FIG. 1 Analysis of the T-cell antigen receptor repertoire during primary HIV infection. A semiquantitative polymerase chain reaction (PCR) assay^{16,17} was used to determine the V β repertoire on PBMC specimens collected at different times after primary HIV infection. Only at least a doubling in the relative expression of V β families among sequential time points by PCR was considered to be significant. METHODS. All patients analysed here presented to hospital emergency rooms and were studied prospectively under protocols approved by the institutional review board of the University of Alabama at Birmingham. Kinetics of viral burden of all six patients have been reported previously along with the clinical histories of patients 1–4 (refs 3–5). Patient 5 was a 17-year-old female who presented with anorexia, nausea, vomiting, mild sore throat and cervical adenopathy. Patient 6 was a 28-year-old homosexual man with a 3 week history of fever, cough, nausea and occasional vomiting. At the time of clinical presentation, enzyme-linked (ELISA) and immunoblot assays for HIV-1 antibody were negative whereas ELISA assays for HIV-1 p24 antigen were positive in both patients. Blood was obtained by venipuncture, and the PBMC populations collected at different times after the onset of symptoms were frozen and stored at -80°C until used. The V β repertoire on unfractionated PBMC was analysed using a semiquantitative PCR assay as

described previously^{16,17}. Briefly, 2 μg of total RNA, extracted by the RNAzol method^{18,19}, were reverse-transcribed using random hexamers (20 $\mu\text{g ml}^{-1}$; Promega), AMV reverse transcriptase (60 U; Life Science) and dNTPs (Boehringer Mannheim). The reaction mixture was incubated at 42°C for 40 min, heated to 94°C for 5 min to inactivate the reverse transcriptase, cooled on ice for 5 min, and then the final reaction volume was diluted to 1:27 with sterile water. PCR analysis for the different V β s used a 5'-specific V β primer and a common 3' constant-domain (C) β primer. 5' and 3' α primers were included in each PCR reaction as internal controls. To monitor the migration of V β and α bands, 3' C β and 3' α primers were radiolabelled with [^{32}P - γ]ATP (Amersham). Primer sequences of the T-cell receptor V β -specific oligonucleotides and of the control C β and α oligonucleotides have been reported¹⁵. Amplification was for 25 cycles in a GenAmp PCR system 9600 (Perkin-Elmer). Half of the volume of each PCR reaction was loaded and separated on 10% polyacrylamide gels containing 7 M urea. Gels were exposed overnight on Kodak storable phosphor screens, and the radioactive signal for each V β was quantified using a PhosphorImager (Molecular Dynamics).

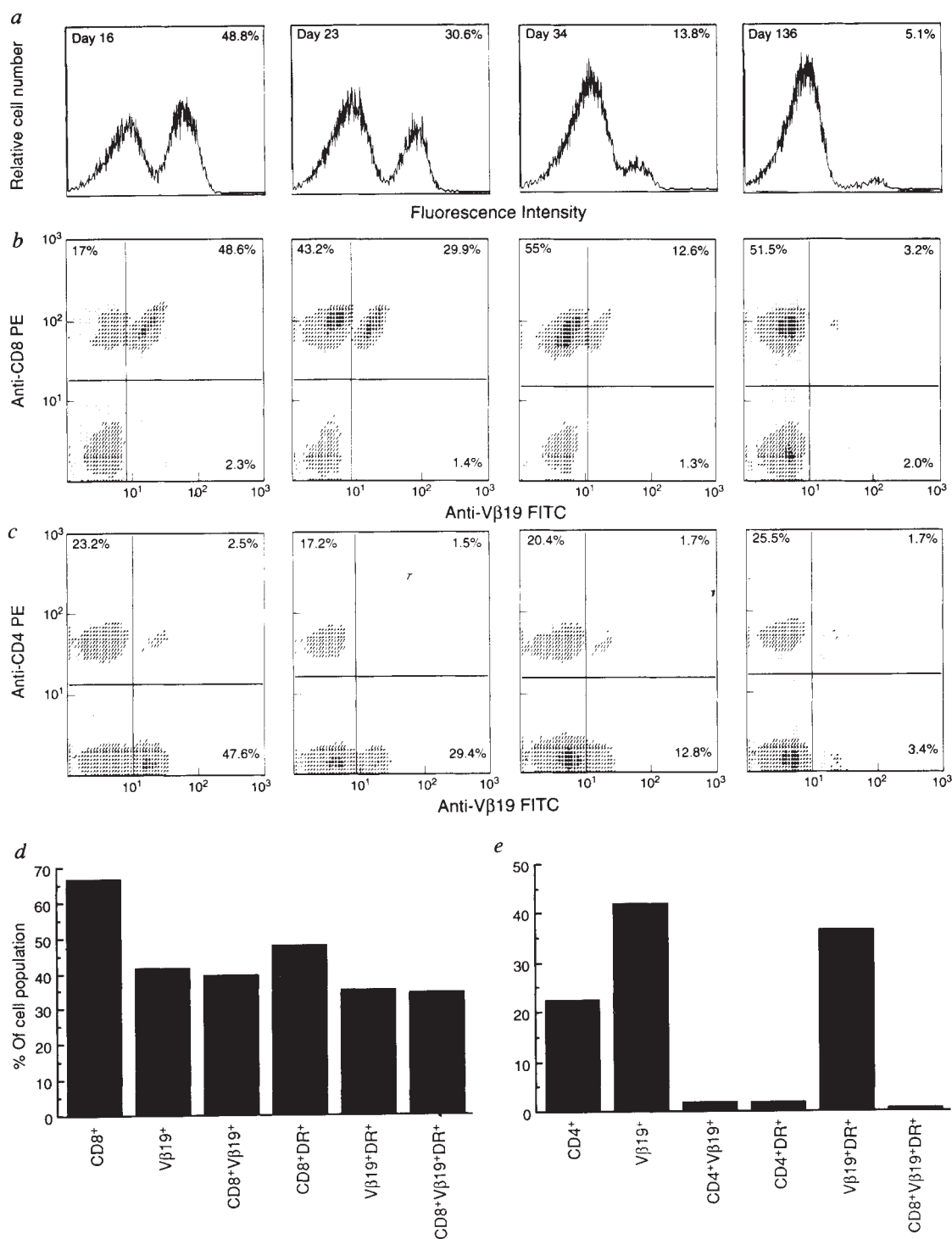


FIG. 2 *a*, Kinetics of Vβ19 expansion in peripheral blood mononuclear cells following primary HIV infection in patient 1. *b*, Two-colour cytofluorometric analysis of CD8 compared with Vβ19. *c*, Two-colour cytofluorometric analysis of CD4 compared with Vβ19. *d*, Three-colour cytofluorometric analysis of Vβ19 compared with CD8 and HLA-DR antigens. *e*, Three-colour cytofluorometric analysis of Vβ19 compared with CD4 and HLA-DR antigens.

METHODS. Two- and three-colour cytofluorometric analyses were done on a Coulter Elite machine as described previously²⁰. Three-colour cytofluorometric analysis was carried out on a PBMC sample collected on day 20 after the onset of symptoms. Cells (5×10^5) were stained simultaneously with FITC (fluorescein isothiocyanate)-conjugated and PE (phycoerythrin)-conjugated monoclonal antibodies (mAbs). Control

aliquots were stained with FITC- or PE-conjugated mouse IgG1 or IgG2. In all experiments, 50,000–60,000 events were accumulated and analysed for fluorescence. FITC-conjugated anti-Vβ19 mAb was from AMAC, and PE-conjugated anti-CD4 and anti-CD8 mAbs were from Becton & Dickinson. For three-colour cytofluorometric analysis, PBMC (1×10^6) were simultaneously stained with FITC-conjugated anti-Vβ19 mAb, PE-conjugated anti-HLA-DR mAb (Becton & Dickinson) and Quantum red-conjugated anti-CD4 or anti-CD8 mAb (Sigma). In all experiments, 50,000–60,000 events were accumulated and analysed for fluorescence. Each possible fluorescent combination was counted using the Prism software. Similar levels of expression of HLA-DR in CD8⁺ and CD4⁺ cell subsets were also observed in patients 2, 3 and 4.

sions strongly suggests that the expanded $V\beta$ s were involved in an antigen-specific immune response.

HIV-specific cytotoxicity tests supported the suggestion that these $V\beta$ expansions resulted from antigen-specific stimulation. PBMC from patient 1 collected on day 16 were stimulated for 10 days with an anti-CD3 monoclonal antibody, and HIV-specific

cytotoxicity was determined against B-lymphoblastoid cell lines transformed with either autologous or allogeneic Epstein-Barr virus (EBV) and infected with vaccinia virus constructs expressing different HIV proteins (HIV *env*-encoded gp160 and a *gag-pol*-encoded fusion protein) and β -galactosidase as control. Unfractionated PBMC mediated MHC-restricted cytotoxic

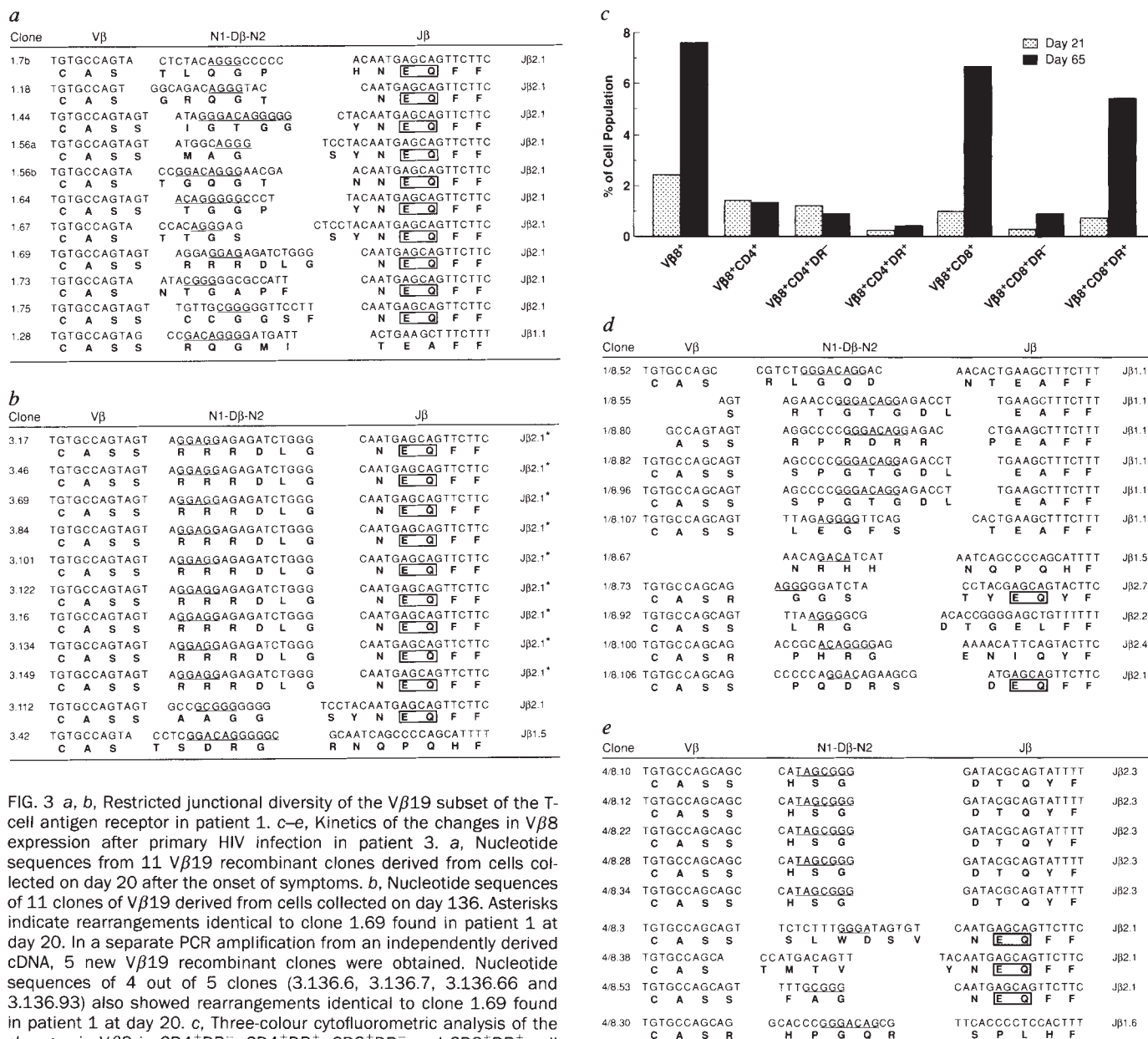


FIG. 3 **a, b**, Restricted junctional diversity of the $V\beta 19$ subset of the T-cell antigen receptor in patient 1. **c–e**, Kinetics of the changes in $V\beta 8$ expression after primary HIV infection in patient 3. **a**, Nucleotide sequences from 11 $V\beta 19$ recombinant clones derived from cells collected on day 20 after the onset of symptoms. **b**, Nucleotide sequences of 11 clones of $V\beta 19$ derived from cells collected on day 136. Asterisks indicate rearrangements identical to clone 1.69 found in patient 1 at day 20. In a separate PCR amplification from an independently derived cDNA, 5 new $V\beta 19$ recombinant clones were obtained. Nucleotide sequences of 4 out of 5 clones (3.136.6, 3.136.7, 3.136.66 and 3.136.93) also showed rearrangements identical to clone 1.69 found in patient 1 at day 20. **c**, Three-colour cytofluorometric analysis of the changes in $V\beta 8$ in $CD4^+DR^-$, $CD4^+DR^+$, $CD8^+DR^-$ and $CD8^+DR^+$ cell populations at days 21 and 65 after the onset of symptoms. The percentages of the other populations analysed were as follows: $V\beta 8^+CD4^+DR^-$ cells were 0.3% and 0.7% at days 21 and 65 respectively; $V\beta 8^+CD4^+DR^+$ were, respectively, 0.7% and 5.7%; $V\beta 8^+CD4^+DR^-$, 30% and 22.2%; $V\beta 8^+CD4^+DR^+$, 44.6% and 45.2%; $V\beta 8^+CD8^+DR^-$, 3.2% and 3.4%; $V\beta 8^+CD8^+DR^+$, 1.2% and 0.9%; $V\beta 8^+CD8^+DR^-$, 0.2% and 0.3%; $V\beta 8^+CD8^+DR^+$, 14.7% and 15.4%; $V\beta 8^+CD8^+DR^-$, 11.7% and 14.6%; $V\beta 8^+CD8^+DR^+$, 32.5% and 33.2%. **d**, Nucleotide sequences of recombinant clones of $V\beta 8$ derived from cells collected on day 6 after the onset of symptoms. **e**, Nucleotide sequences of recombinant clones of $V\beta 8$ derived from cells collected on day 65. V, variable; N, non-templated; D, diversity; J, joining.

METHODS. Total RNA was extracted from PBMC of patients 1 and 3 collected at different times after the onset of symptoms, and complementary DNA synthesis was done as described in Fig. 1 legend. To enrich for $V\beta 19$ and $V\beta 8$ cDNAs, total cDNA was amplified for 35 cycles

in a GenAmp PCR system 9600. Primers used were HBVT-02 *Sall* for $V\beta 19$ (5'-AAGGTCGACATGAGCAACCAGGTGCTCTGCTGT3') and $V\beta 8$ *Sall* (5'-AACCGTCGACTCTGGTACAGACAGACCATGAT3'), and a common 3' *C β F* (5'-GGTGTGGGAGATGCTGACTTTTGATGGCTCAAAC3'). After overnight digestion with *Sall* (20 U) the amplification product was cloned into the *Sall*-digested dephosphorylated pYS eukaryotic expression vector. Nucleotide sequence of the recombinant clones was determined by the dideoxy termination sequencing method using the Pharmacia T7 sequencing kit. Primers used were mC β V (5'-AGGATTGTGCCAGAAGGTAGC3') and RSVneoLTR (5'-GTGCCTTATTAGGAAGCAAC3'). Sequences were aligned to the published HBVT02 (ref. 21) ($V\beta 19$), $V\beta 8$ (ref. 22) and D β (ref. 14) J β sequences⁶. The regions of homology to the germline sequence are underlined. Conserved EQ residues in the J β region are boxed. Three-colour cytofluorometric analysis was as described in Fig. 2. Legend. FITC-conjugated anti- $V\beta 8$ mAb was from AMAC.

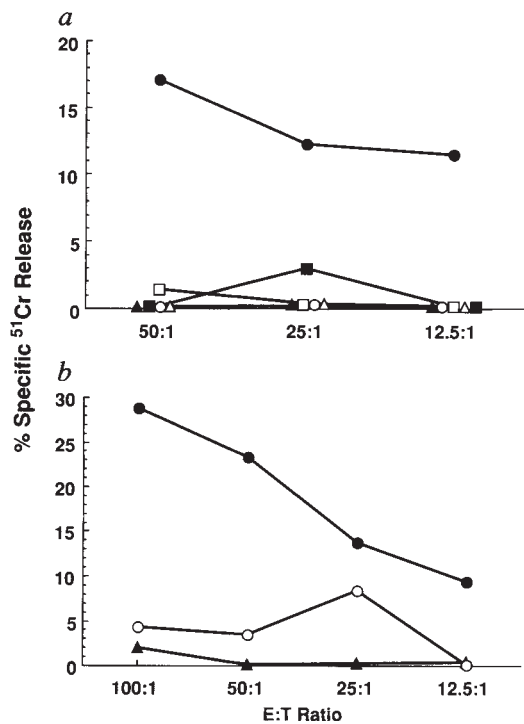


FIG. 4 HIV-specific cytotoxic activity in unfractionated PBMC and sorted Vβ19⁺CD8⁺ T cells from patient 1. a, Unfractionated PBMC; b, sorted Vβ19⁺CD8⁺ T cells. ▲ β-galactosidase; △ allogeneic β-galactosidase; ●, HIV env; ■, allogeneic HIV env; ○ HIV gag-pol; □ allogeneic HIV gag-pol.

METHODS. A frozen sample of PBMC collected on day 16 after the onset of symptoms was thawed and cells were stimulated using an anti-CD3 mAb as described¹⁵. Briefly, patient PBMC (5×10^5 cells per well) were resuspended in RPMI 1640 supplemented with 10% fetal bovine serum and 100 U ml^{-1} recombinant interleukin-2 (Boehringer Mannheim) in 24-well plates in the presence of $0.4 \mu\text{g ml}^{-1}$ anti-CD3 mAb (Coulter Immunology) and 5×10^6 γ-irradiated (3,500 rad) allogeneic PBMC per well. After 10 days of culture the proliferating cell population was tested for HIV-specific cytotoxic activity. The remaining cells were further expanded for an additional week and then sorted for CD8⁺Vβ19⁺ cells by two-colour sorting as described²³. The purity of the sorted population was 98%. Cytotoxic activity was assayed in unfractionated and sorted CD8⁺Vβ19⁺ cell populations in a conventional 4-h ⁵¹Cr-release assay as described²³. Autologous and allogeneic EBV B-lymphoblastoid cell lines were infected at a multiplicity of infection of 5–10 plaque-forming units per cell for 12–16 h with vaccinia constructs expressing different HIV proteins and used as target cells. The MHC class I typing of patient 1 was A1, 29; B8, 44, whereas HLA A3, 19; B18, 35-positive EBV-transformed B-lymphoblastoid cell lines were used as allogeneic targets. Cytotoxic activity against allogeneic target cells infected with vaccinia virus constructs expressing HIV proteins was determined only when unfractionated PBMC were used as effector cells. Recombinant vaccinia viruses used here included vPE16, which expresses gp 160 from the BH8 clone of HIV-1 IIB²⁴, vVKI, which expresses a gag-pol fusion protein²⁵, and vSC8 (control), which expresses only *Escherichia coli* β-galactosidase²⁶. These vaccinia constructs were provided by B. Moss (Laboratory of Molecular Virology NIAID, NIH). Labelling of target cells with Na⁵¹CrO₄ (Amersham) and the cytotoxic assays were carried out as described²³. The per cent specific ⁵¹Cr release was calculated as: $100 \times (\text{experimental release} - \text{spontaneous release}) / (\text{total release} - \text{spontaneous release})$. Results are expressed as % specific ⁵¹Cr release.

activity against autologous EBV B-lymphoblastoid cell lines expressing HIV env. (Fig. 4a). Similar cytotoxic T-lymphocyte activity in unfractionated PBMC was found in patients 2, 3 and 4 (ref. 8). However, sorted CD8⁺Vβ19⁺ T cells mediated HIV-specific cytotoxicity against autologous target cells expressing HIV env but not against cells expressing gag-pol or the β-galactosidase control (Fig. 4b). The presence of HIV-specific cytotoxic T lymphocytes in the CD8⁺Vβ19⁺ T-cell subset indicates that their expansion was driven by HIV antigens.

The fact that a high effector-to-target cell ratio was needed for significant levels of specific lysis, despite the use of an *in vitro*-activated, purified Vβ19⁺CD8⁺ cell population, may suggest differences in the epitopes of the env expressed by the vaccinia virus and those expressed in the patient; alternatively, these CD8⁺ T cells may be poorly cytotoxic because of their dramatic expansion.

Neither of two previous studies reporting expansion of Vβ families^{9,10} gave direct evidence for the role of selective expan-

sions of Vβ subsets of CD8⁺ T lymphocytes in the specific immune response to microbial pathogens. Here we provide both functional and molecular evidence that these oligoclonal Vβ expansions are the result of a specific immune response to HIV. The biological and clinical significance of this type of immune response is unknown. Notably, however, patients 1 and 5, who experienced the most dramatic Vβ expansion, progressed to AIDS within 12 months from the time of seroconversion. These Vβ expansions may reflect the presence of a limited T-cell receptor repertoire in HIV infection¹¹ which facilitates virus escape from the immune response^{11,12} by limiting recognition of virus variants. Alternatively major Vβ expansions may be associated with an exhaustion of antiviral cytotoxic effector cells resulting in the escape of the virus from elimination by the immune system¹³. This type of immune response may be typical of other viral infections and may be critical for antiviral immune surveillance. □

Received 4 May; accepted 27 June 1994.

- Pantaleo, G., Graziosi, C. & Fauci, A. S. *New Engl. J. Med.* **328**, 327–335 (1993).
- Daar, E. S., Moudgil, T., Meyer, R. D. & Ho, D. D. *New Engl. J. Med.* **324**, 961–964 (1991).
- Clark, S. J. et al. *New Engl. J. Med.* **324**, 954–960 (1991).
- Graziosi, C., et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 6405–6409 (1993).
- Piatk, M. et al. *Science* **259**, 1749–1754 (1993).
- Toyonaga, B., Yoshikai, Y., Vadasz, V., Chin, B. & Mak, T. W. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8624–8628 (1985).
- Jorgensen, J. L., Reay, P. A., Ehrlich, E. W. & Davis, M. A. *Rev. Immun.* **10**, 835–873 (1992).
- Borrow, P., Lewicki, H., Hahn, B. H., Shaw, G. M. & Oldstone, M. B. A. *J. Virol.* (In the press).
- MacDonald, H. R., Casanova, J.-L., Maryanski, J. L. & Cerottini, J.-C. *J. exp. Med.* **177**, 1487–1492 (1993).
- Smith, T.-J., Terada, N., Robinson, C. C. & Gelfand, E. W. *Blood* **81**, 1521–1526 (1993).
- Kalams, S. A. et al. *J. exp. Med.* **179**, 1261–1271 (1994).
- Phillips, R. E. et al. *Nature* **354**, 453–459 (1991).
- Moskophidis, D., Lechner, F., Pircher, H. & Zinkernagel, R. M. *Nature* **362**, 758–761 (1993).
- Ferradini, L. et al. *Eur. J. Immun.* **4**, 935–942 (1991).
- Walker, et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9514–9518 (1989).
- Labrecque, N., McGrath, M., Subramanyam, M., Huber, B. T. & Sekaly, R.-P. *J. exp. Med.* **177**, 1735–1743 (1993).
- Rebai, N. et al. *Proc. natn. Acad. Sci. U.S.A.* **91**, 1529–1533 (1994).

- Chomczynski, P. & Sacchi, N. *Analyt. Biochem.* **162**, 156–159 (1987).
- Pantaleo, G. et al. *Nature* **362**, 355–358 (1993).
- Pantaleo, G., Koenig, S., Baseler, M., Lane, H. C. & Fauci, A. S. *J. Immun.* **144**, 1696–1704 (1990).
- Kimura, N., Toyonaga, B., Yoshikai, Y., Du, R. P. & Mak, T. W. *Eur. J. Immun.* **17**, 375–383 (1987).
- Tillinghast, J. P., Behlke, M. A. & Loh, D. Y. *Science* **233**, 879–883 (1986).
- Pantaleo, G. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4818–4822 (1990).
- Earl, P. L., Hugin, A. & Moss, B. *J. Virol.* **64**, 2448–2451 (1990).
- Karacostas, V., Nagashima, K., Gonola, M. A. & Moss, B. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8964–8967 (1989).
- Chakrabarti, S., Brechling, K. & Moss, B. *Molec. cell. Biol.* **5**, 3403–3409 (1985).

ACKNOWLEDGEMENTS. We thank H. R. MacDonald, J. C. Cerottini and B. D. Walker for suggestions and discussion, M. B. A. Oldstone for sharing unpublished results, and M. A. Rust for editorial assistance. This work counts in partial fulfilment of the PhD requirements for J.F.D. in the Graduate Genetics Program at George Washington University. Clinical specimens were obtained with signed informed consent of the study subjects and were processed and stored in the care support facilities of the University of Alabama Center for AIDS Research. Supported in part by the NIMH AIDS center MRC and the NHRDP-AIDS Program. R.P.S. holds a MRC scientist award. F.D. is a postdoctoral fellow of NHRDP.

Implications for dorsoventral axis determination from the zebrafish mutation *janus*

Salim Abdelilah, Lilianna Solnica-Krezel, Didier Y. R. Stainier & Wolfgang Driever

Cardiovascular Research Center, Massachusetts General Hospital and Harvard Medical School, 13th Street, Bldg 149, Charlestown Massachusetts 02129, USA

THE mechanisms underlying the formation of dorsoventral polarity in the zebrafish *Danio rerio* are unknown. Here we describe the zebrafish recessive maternal-effect mutation *janus*^{ms5}. The mutant phenotype is a division of the blastoderm along the first cleavage plane into two detached half-sized blastoderms. Partial-axis bifurcation occurs in a subset of mutants. Analysis of *goosecoid* expression in the mutant embryos indicates that only one organizer region is present in each embryo. Furthermore, the position of this organizer region is random with respect to the first cleavage plane bisecting the two blastoderms. Finally, cell tracing in wild-type embryos demonstrates that there is no strict correlation of the dorsoventral axis with early cleavage planes in zebrafish. These findings support the notion that the establishment of the dorsoventral axis and the first cleavage planes are determined by separate mechanisms in the zebrafish embryo.

The mutation *janus*^{ms5} (*jan*^{ms5}) is a spontaneous, temperature-sensitive, strictly maternal and recessive mutation. The mutant phenotype becomes detectable during the first three cleavages, which coincides with the temperature-sensitive period of the mutation (data not shown). In contrast to wild-type, mutant blastomeres clearly separate along the first cleavage plane (17/17 embryos analysed). This spatial separation results in a division of the early blastoderm into two groups of blastomeres by the 16-cell stage (Fig. 1a). We found the penetrance of this early phenotype to be highly variable, ranging from 5–90%, depending mainly on the temperature but also on the genetic background. The two resulting blastoderms continue their early development

separated from one another unless they fuse (S.A. and W.D., manuscript in preparation).

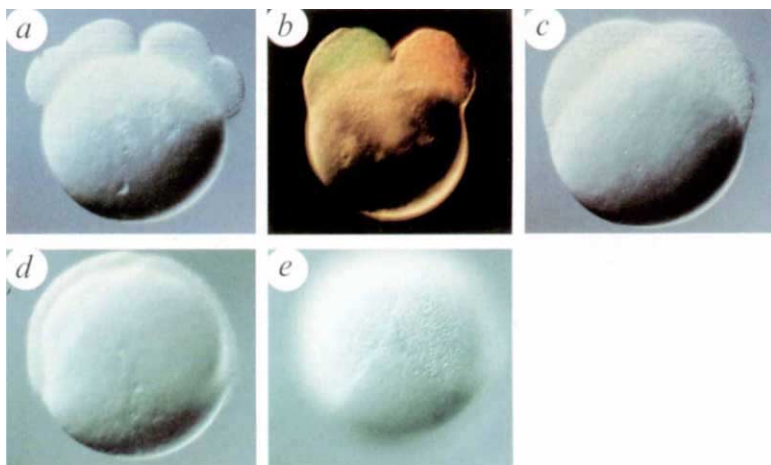
To test whether cell mixing occurs between the blastoderms, we did a cell tracing experiment in which we injected two blastomeres at the 2-, 4- or 8-cell stage on opposite sides of the first cleavage plane with different fluorescently labelled dextrans¹. All 17 injected mutant embryos showed two distinctly labelled blastoderms during blastula stages (Fig. 1b). No cell mixing occurred between the two blastoderms before the onset of epiboly.

During epiboly in teleost embryos, the blastoderm spreads toward the vegetal pole. A dorsoventral asymmetry becomes apparent when the blastoderm covers 50% of the yolk cell. Gastrulation movements create the shield, a cellular thickening on the dorsal side of the embryo². In mutant embryos, the shield is formed when each blastoderm covers ~25–30% of the yolk (Fig. 1d, e). Epibolic movements cause the blastoderms to come into contact with one another. In progeny of homozygous *jan*^{ms5} females, among those embryos that have double blastoderms, only about 25% develop axial bifurcations to a variable extent (Fig. 2). About 34% (16/47) of these abnormal embryos have symmetrical duplications of anterior structures (Fig. 2a); the other 66% have additions of embryonic tissues along the trunk or tail regions (Fig. 2b).

Double blastoderms have been reported in zebrafish and other teleosts^{3,4}. It was suggested that axis bifurcations result from duplications of the organizer region⁴. To determine whether more than one organizer region⁵ is present in *jan*^{ms5} embryos, we used *goosecoid* as an early marker for prospective dorsal mesoderm^{6,7}. We examined the position of the dorsal organizer region relative to the plane bisecting mutant blastoderms and detected only one organizer region in each mutant embryo. Its location is random with respect to the plane that divides the two blastoderms (Fig. 2d–f). Thus, in two thirds of the *jan*^{ms5} embryos, only one of the two blastoderms has an organizer region. In the remaining third, an organizer region lies within the area where the division of the two blastoderms occurred (Fig. 2d, e). In some of these cases the organizer is clearly partitioned between the two blastoderms. Among the mutant embryos with a partitioned organizer, about half (14/25) have an organizer

FIG. 1 Early development of *janus* mutant embryos at 32°C. a, By the 16-cell-stage the blastoderm is divided into two groups of cells. b, A mutant 128-cell stage embryo that was double-injected at the two-cell stage exhibits two uniquely labelled blastoderms. No cell-mixing between the blastoderms occurred upon separation. c, Late blastula; d, and e, at the shield stage, each blastoderm covers 25–30% of the yolk: d, deep and e, shallow plane of focus. The mutation *jan*^{ms5} was found as a naturally occurring mutation in a zebrafish strain that was obtained from Hong Kong. Genetic test crosses show that the mutation is maternal and has a recessive mode of segregation (S.A. and W.D., manuscript in preparation). Therefore, it is one of the few known maternal-effect mutations in vertebrates^{16–20}.

METHODS. Embryos were embedded in 1.5% methylcellulose in 10% Hank's balanced salt solution (HBSS) for photographic documentation. For injection, embryos were collected immediately upon egg laying and kept at 32 °C since *jan*^{ms5} is temperature-sensitive with the highest penetrance of separated blastoderms at this temperature (S.A. and W.D., manuscript in preparation). Wild-type embryos develop normally at temperatures between 23–34 °C (ref. 21). After manual dechoriation, embryos were double-injected with a 50 mg ml⁻¹ solution of high-molecular-weight (2,000 K) fluorescein (Sigma) or Texas red (Molecular Probes) conjugated uncharged dextran in 0.1 M KCl. High molecular weight lineage label was used to minimize leakage of dye



into neighbouring cells at stages when blastoderm cells are still not completely separated from one another by membranes¹. A total of 17 mutant embryos of either the 2-, 4- or 8-cell stage were double-injected into blastomeres on opposite sides of the first cleavage plane by pressure driven microinjection.

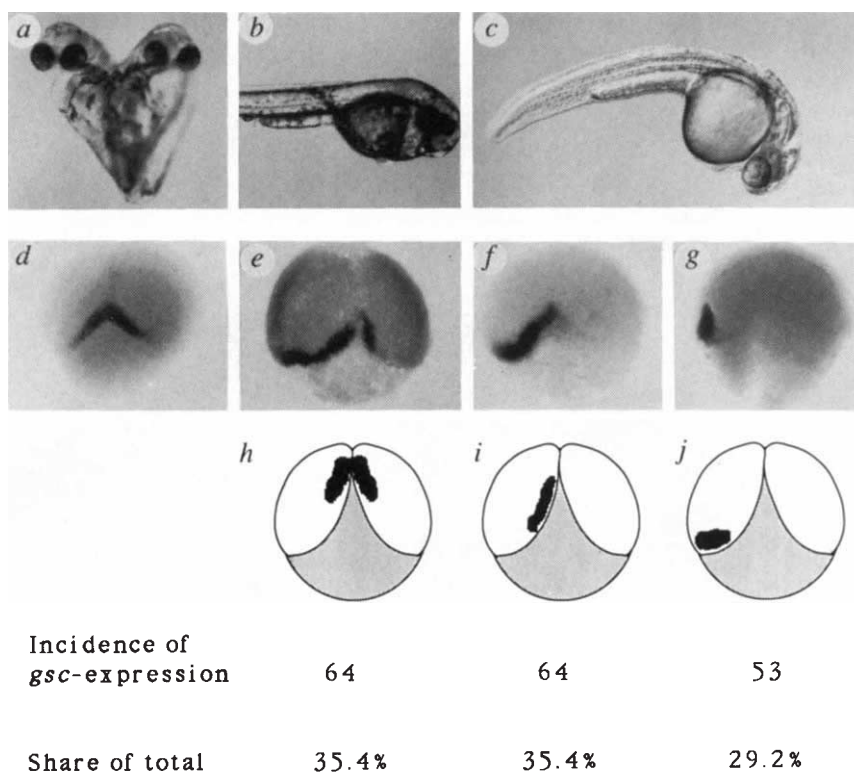


FIG. 2 Examples of *janus*-phenotypes after 24 h of development and incidence of different *goosecoid*-expression patterns of pregastrulation- and gastrulation-stage *janus* mutant embryos. About 34% of the axis bifurcations are symmetrical duplications of an anterior region (a); the rest have additions of embryonic tissues mainly in their trunk or tail regions (b). c, A phenotypically normal embryo after 24 h. A total of 1,649 fertilized eggs resulted in 319 (19%) phenotypically mutant embryos at the shield stage, 164 of which were allowed to develop further. After 24 h, 47 mutant axis formations were scored: 16 (34%) show symmetrical bifurcations of anterior regions, 31 (66%) display ectopic tissue formation. The *gsc*-expression was analysed in a total of 181 embryos. The dorsal organizer region correlates either with the first

cleavage plane bisecting both blastoderms (d, e, h), an intermediate plane between the first and second cleavage plane (f, i), or with the second cleavage plane (g, j). Among embryos that show a correlation between the organizer region and the first cleavage plane (d, e), a distinction between equally and unequally distributed organizer regions could be made for a subset of 25 pre-shield stage embryos: d, Organizer region is equally distributed between the two blastoderms (14/25); e, Organizer region is unequally distributed between the two blastoderms (11/25). We suggest a correlation of the later phenotypes with the different patterns of distribution of the *gsc*-expressing region as shown in this figure. *In situ* hybridization was done as previously described²².

equally distributed between the blastoderms (Fig. 2d) whereas in the other half (11/25) the organizer is unequally distributed (Fig. 2e). We suggest that symmetrical axis duplications of the anterior region occur in embryos that have an organizer equally distributed between the two blastoderms. Ectopic additions of embryonic tissues might result from an unequally distributed organizer.

The spatial randomness of the organizer region with respect to the first cleavage plane indicates that the dorsoventral axis in *jan^{ms5}* mutant embryos does not coincide with the second cleavage plane. Contrary to this finding is a report claiming that the dorsoventral axis in wild type strictly correlates with the second cleavage plane¹. It is therefore possible that the indeterminacy of the dorsoventral axis with respect to a specific cleavage plane in *jan^{ms5}* mutant embryos could be part of the mutant phenotype. To test this, we did a cell-tracing experiment in wild-type embryos similar to one recently reported¹. Two blastomeres of the 8-cell-stage embryo were injected with fluorescently labelled dextrans (Fig. 3a). The position of the clonal progeny of these cells with respect to the dorsoventral axis was analysed at the shield stage (Fig. 3f-j) and again during early somitogenesis (Fig. 3k-y). An advantage of scoring at the shield stage compared with later stages, as was previously described¹, is that

dorsal and ventral cell groups are well separated and can be easily distinguished (Fig. 3f-j)⁸. The subsequent convergence movements bring dorsal and ventral cells into close proximity within the embryonic body. Unambiguous classification of the fates of labelled cells as dorsal or ventral is thus more difficult in older embryos.

Our results show no strict correlation of any cleavage plane with the dorsoventral axis (Fig. 3). The position of the dorsoventral axis was scored as coinciding with either the first ($\pm 30^\circ$) or second ($\pm 30^\circ$) cleavage plane, or with neither of them (Fig. 3 legend). Hence, we expect 33% incidence in each of these three categories by chance alone. The observed coincidence of second cleavage plane with later dorsoventral axis (43.2%) is only slightly biased (Fig. 3g, l, q, v). Similar results favouring a certain bias over total randomness have been reported using cell tracing in embryos at later stages or in different teleost species⁸⁻¹¹.

In the amphibian *Xenopus laevis*, the dorsoventral polarity is established by cortical rotation of the egg cytoplasm following fertilization^{12,13}. The first cleavage planes and the embryonic axes appear to be determined by separate mechanisms^{14,15}. The work described here indicates that in zebrafish, cleavage planes cannot be used to predict the position of the future dorsoventral axis.

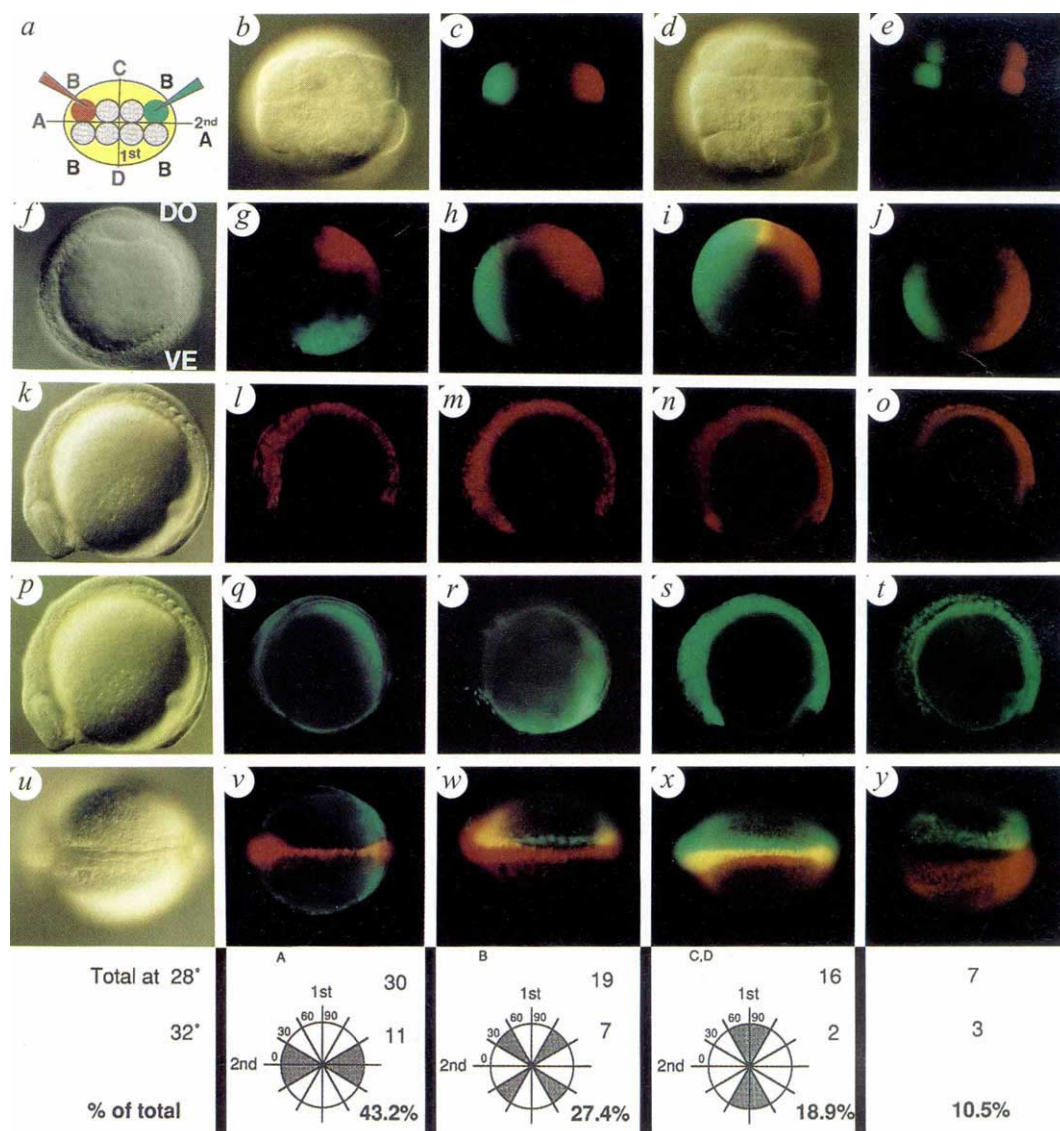


FIG. 3 Cell-tracing study in wild-type embryos. *a*, Two blastomeres at opposite ends of a 4-cell row in an 8-cell stage blastoderm were labelled green (fluorescein) or red (Texas red). The letters A, B, C, D indicate positions where the dorsal side was found to form with respect to the early cleavage planes. In each vertical row (*f*–*y*), micrographs are of one single experimental embryo at different developmental stages. *g*, *l*, *q*, *v*, example for A: dorsoventral axis correlates with second cleavage plane; *h*, *m*, *r*, *w*, example for B: dorsoventral axis correlates with intermediate plane between first and second cleavage plane; *i*, *n*, *s*, *x*/*j*, *o*, *t*, *y*, examples for C/D: dorsoventral axis correlates with first cleavage plane; *i*, *n*, *s*, *x*, example for C: dorsal side will form 'ipsilateral' to injected cells; *j*, *o*, *t*, *y*, example for D: dorsal side will form 'contralateral' to injected cells. Stage of embryos: *b*, *c*, 8-cell stage; *d*, *e*, 16-cell stage; *f*–*j*, shield stage—gastrulation movements create a dorsal thickening; *k*–*y*, 8-somite- to 12-somite-stage embryos. Orientation: *b*–*j*, animal view; *f*–*j*, dorsal at top; *k*–*t*, lateral view with anterior left, dorsal at top; *u*–*y*, dorsal view, anterior left. At the bottom, the numbers of embryos that developed according to the categories displayed in the vertical rows are given for two different experimental temperatures. Categories were assigned according to distribution of clones at the shield stage. At later stages, distribution of clones within the adult body was found to correlate with the predicted positions according to a fate map for the gastrulation-stage zebrafish embryo²³.

METHODS. The experiment used zebrafish 'AB' wild-type strain (Univ. of Oregon, Eugene). Injections were done as for Fig. 1. Embryos were collected immediately upon egg-lay and manually dechorionated before

the 4-cell stage. Another sample group of 17 embryos was injected in their chorions, which were manually removed during late-blastula stages. Injections were done on embryos kept at 28 °C. Other embryos ($n=23$) were also injected at 32 °C, the temperature at which *jan*^{m55} mutant embryos are kept. The different treatment or temperature had no significant effect on the correlation of early cleavage planes with the later dorsoventral axis. A total of 410 embryos were double-injected, 95 of which were counted as 'well-injected'. Criteria for 'well-injected' embryos were as follows: pattern of the early cleavage stages was not perturbed; injected cells faithfully transmitted the dye into small clones of 2, 4 etc. descendant cells according to the stage of development; embryos survived and appeared normal at the early-somite stages. A random sample of 15 double-injected embryos was incubated further; all survived for longer than 36 h. The position of the dorsoventral axis with respect to the cleavage planes was analysed at the shield stage. Three categories were distinguished: A, correlation of dorsoventral axis with the second cleavage plane was assumed for cases in which the shield formed $\pm 30^\circ$ from the second cleavage plane that transects the labelled clones; C, D, correlation of the dorsoventral axis with the first cleavage plane was assumed for cases in which the shield formed $\pm 30^\circ$ from the first cleavage plane, the plane perpendicular to the plane that transects the two labelled clones; B, lack of correlation of the dorsoventral axis with the first two division planes was assumed for embryos in which the shield was positioned at an angle larger than 30° from the two cleavage planes.

Therefore, we suggest that in zebrafish as in *Xenopus laevis*, the first cleavage plane and the dorsoventral axis are determined by separate mechanisms. □

Received 27 April; accepted 14 July 1994.

1. Strehlow, D. & Gilbert, W. *Nature* **361**, 451–453 (1993).
2. Westerfield, M. *The Zebrafish Book* (Univ. Oregon Press, USA, 1993).
3. Laale, H. W. *Can. J. Zool.* **62**, 386–390 (1984).
4. Carter, C. A. & Wourms, J. P. *J. Morph.* **215**, 301–312 (1993).
5. Stachel, S. E., Grunwald, D. J. & Myers, P. Z. *Development* **117**, 1261–1274 (1993).
6. Blumberg, B., Wright, C. V. E., DeRobertis, E. M. & Cho, K. W. Y. *Science* **253**, 194–196 (1991).
7. Cho, K. W. Y., Blumberg, B., Steinbeisser, H. & DeRobertis, E. M. *Cell* **67**, 1111–1120 (1991).
8. Kimmel, C. B. & Warga, R. M. *Dev. Biol.* **124**, 269–280 (1987).
9. Oppenheimer, J. M. *J. exp. Zool.* **73**, 405–444 (1936).
10. Kimmel, C. B. & Law, R. D. *Dev. Biol.* **108**, 78–85 (1985).
11. Gevers, P., Dulos, J., van den Boogaart, J. G. M. & Timmermans, L. P. M. *Roux's Arch. dev. Biol.* **201**, 275–283 (1992).
12. Gerhart, J., Ubbels, G., Black, S., Hara, K. & Kirschner, M. *Nature* **292**, 511–516 (1981).
13. Gerhart, J. et al. *Development* suppl. **107**, 37–51 (1989).
14. Danilchik, M. V. & Black, S. D. *Dev. Biol.* **128**, 58–64 (1988).
15. Black, S. D. & Vincent, J. P. *Dev. Biol.* **128**, 65–71 (1988).
16. Carrol, C. R. & Van Deusen, E. B. *Dev. Biol.* **32**, 155–166 (1973).
17. Malacinski, G. M. & Brothers, A. J. *Science* **184**, 1142–1147 (1974).
18. Droin, A. & Fischberg, M. *Roux's Arch. dev. Biol.* **193**, 86–89 (1984).
19. Ikenishi, K. & Tsuzaki, Y. *Dev. Biol.* **125**, 458–461 (1988).
20. Kubota, H. Y., Itoh, K. & Asada-Kubota, M. *Dev. Biol.* **144**, 145–151 (1991).
21. Schirone, R. C. & Gross, L. J. *exp. Embryol.* **169**, 43–52 (1968).
22. Oxtoby, E. & Trevarrow, J. *Nucleic Acids Res.* **21**, 1087–1095 (1993).
23. Kimmel, C. B., Warga, R. M. & Schilling, T. F. *Development* **108**, 581–594 (1990).
24. Helde, K. A., Wilson, E. T., Cretekos, C. J. & Grunwald, D. J. *Science* **265**, 517–520 (1994).

ACKNOWLEDGEMENTS. We thank M. Fishman for discussion and support; D. Stemple, S. Neuhauss, A. Schier, F. Zwartkruis, D. Kane, M. Mullins, N. Hopkins, D. Strehlow, W. Gilbert and P. Cohen for comments on an earlier version of the manuscript; and S. Schulte-Merker for the *goosecoid* probe. Another investigation of the relation of cleavage planes versus axis formation has led to similar conclusions²⁴. This work was supported by a grant from the NIH and by Bristol-Myers Squibb. D.Y.R.S. is supported by the Helen Hay Whitney Foundation.

Cell-free formation of protease-resistant prion protein

David A. Kocisko^{*†}, Jon H. Come^{*†},
Suzette A. Priola[†], Bruce Chesebro[†],
Gregory J. Raymond[†], Peter T. Lansbury^{*†}
& Byron Caughey^{*†}

^{*} Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

[†] Laboratory of Persistent Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Rocky Mountain Laboratories, Hamilton, Montana 59840, USA

THE infectious agent (or 'prion') of the transmissible spongiform encephalopathies (TSEs) such as scrapie resembles a virus in that it replicates *in vivo* and has distinct strains¹, but it was postulated long ago to contain only protein^{2,3}. More recently, PrP^{Sc}, a pathogenic, scrapie-associated form of the host-encoded prion protein (PrP), was identified as a possible primary TSE agent protein^{4–6}. PrP^{Sc} is defined biochemically by its insolubility and resistance to proteases⁷ and is derived post-translationally from normal, protease-sensitive PrP (PrP^C)^{8,9}. The conversion seems to involve conformational change rather than covalent modification^{10–13}. However, the conversion mechanism and the relationship of PrP^{Sc} formation to TSE agent replication remain unclear. Here we report the conversion of PrP^C to protease-resistant forms similar to PrP^{Sc} in a cell-free system composed of substantially purified constituents. This conversion was selective and required the presence of preexisting PrP^{Sc}, providing direct evidence that PrP^{Sc} derives from specific PrP^C–PrP^{Sc} interactions.

† To whom correspondence should be addressed.

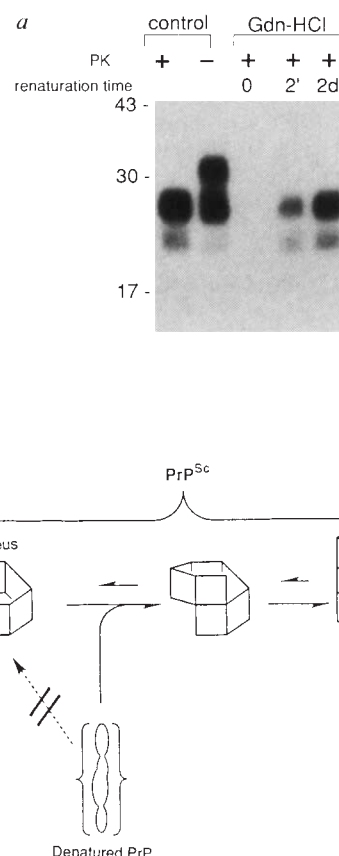


FIG. 1 a, Reversible denaturation of PrP^{Sc} with GdnHCl. Immunoblot analysis of the effect of 3 M GdnHCl treatment on the PK resistance of PrP^{Sc}. As an indicator of the native state of PrP^{Sc}, we used the characteristic partial proteinase K (PK) resistance which distinguishes PrP^{Sc} from PrP^C (ref. 7) and most other proteins (as shown in Fig. 3). Digestion of hamster PrP^{Sc} with PK removes ~67 amino-acid residues (6–7K) from the amino terminus but leaves a large carboxy-terminal fragment (~143 residues) of the molecule intact^{10,27} (compare the first and second lanes). In contrast, PrP^C (ref. 7) and GdnHCl-treated PrP^{Sc} (third lane) are much more completely digested by PK. The antiserum used for the detection of PrP was directed against a synthetic peptide corresponding to hamster PrP residues 90–104 (ref. 28). This antiserum recognizes three-differentially glycosylated forms of PrP with or without PK treatment. 'Renaturation time' is the time in minutes (') or days (d) between the dilution of the 3 M GdnHCl and the start of PK digestion. The positions of molecular mass markers are shown on the left ($\times 10^{-3}$). b, A possible scenario to explain the observed reversibility of PrP^{Sc} denaturation. PrP^{Sc} is shown as an aggregate comprising a conformational isoform (square) of PrP^C. The formation of PrP^{Sc} is predicted to be rate-limited by nucleus formation and hence will probably not occur in the absence of a nucleus, or seed (depicted as a six-mer)^{19,25}. Renaturation of PrP^{Sc} is observed only when some PrP^{Sc} remains after the denaturation step to act as an *in situ* seed. When all the PrP^{Sc} is denatured and disaggregated, for example after treatment with 6 M GdnHCl, renaturation of PrP^{Sc} does not occur because nucleus formation is not favoured under these conditions (dashed arrow). METHODS. PrP^{Sc} was purified from scrapie-infected (263K strain) hamster brain as described previously¹¹ except that the final proteinase K digestion of the PrP^{Sc} (P₄) was omitted. PrP^{Sc} was incubated for 8–24 h at $1 \mu\text{g } \mu\text{l}^{-1}$ in 3 M GdnHCl. One aliquot was treated for 1 h with PK in 3 M GdnHCl (renaturation time, 0). Equivalent aliquots were diluted to 0.75 M GdnHCl with 130 mM NaCl, 10 mM Tris-HCl pH 7.0 (TN) and incubated for ~2 min or 2 days before further dilution to 0.075 M GdnHCl with TN and PK treatment. PK was added to give a final concentration of $50 \mu\text{g } \text{ml}^{-1}$ and the sample was incubated for 1 h. All treatments and incubations of PrP^{Sc} in GdnHCl and PK were at 37 °C. The samples were treated with a PK inhibitor (Pefabloc) and SDS-PAGE sample buffer and immunoblotted²³ with the same equivalents of the original PrP^{Sc} loaded into each lane.

Our strategy for simplifying the analysis of PrP^{Sc} formation was to develop conditions which allowed the reversible denaturation of PrP^{Sc} and then to test if *de novo* conversion of PrP^C into a proteinase K (PK)-resistant form could be observed under these conditions. PrP^{Sc} was denatured with the chaotropic salt guanidine-HCl¹⁴ (GdnHCl), which greatly enhanced its PK sensitivity (Fig. 1). After dilution of the GdnHCl and incubation for two days, most of the PrP regained its original level of PK resistance, indicating that the denaturation in 3 M GdnHCl at 37 °C was reversible. Under harsher conditions (≥ 3.5 M GdnHCl at 37 °C, 3 M GdnHCl at 50 °C or ≥ 2 M guanidine thiocyanate) the denaturation was not reversible (data not shown), which suggested that the denaturation induced by 3 M GdnHCl at 37 °C was incomplete and that reversibility of the denaturation required the maintenance of some native PrP^{Sc} structure.

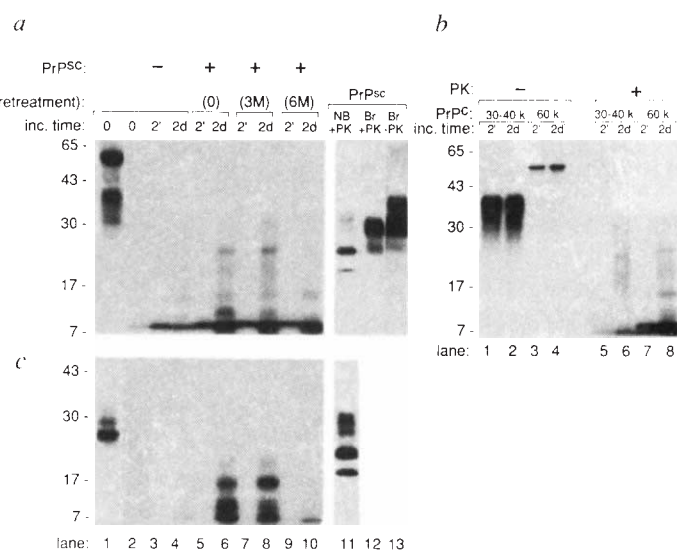
To evaluate whether PrP^C could be converted to a PK-resistant species during the PrP^{Sc} renaturation reaction, ³⁵S-labelled recombinant hamster PrP^C derived from uninfected tissue culture cells was added. After incubation for two days with PrP^{Sc}, PK digestion eliminated the original full-length ³⁵S-PrP^C and produced ³⁵S-labelled PK-resistant PrP fragments of comparable size to PK-digested PrP^{Sc} from scrapie-infected sources (Fig. 2). These PK-resistant ³⁵S-PrP bands were not observed in the absence of PrP^{Sc} and were greatly reduced in the presence of PrP^{Sc} pretreated with 6 M GdnHCl, indicating that some native

PrP^{Sc} structure was required for the formation of large PK-resistant species. Partial denaturation of PrP^{Sc} with 3 M GdnHCl improved the efficiency of the conversions, but was not essential. Preservation of the native conformation of the ³⁵S-PrP^C species did not seem to be required because these species converted to PK-resistant forms after pretreatment with either 3 M GdnHCl (Fig. 2) or 6 M GdnHCl (data not shown).

Because part of the complexity of the pattern of ³⁵S-labelled PK-resistant PrP bands generated from the full-length hamster PrP gene (Fig. 2*a, b*) was probably due to the differential *N*-linked glycosylation of the PrP^C precursor^{15,16}, we carried out the conversion reaction using PrP^C labelled in the presence of tunicamycin, an inhibitor of *N*-linked glycosylation. A prominent PK-resistant PrP species of relative molecular mass 19,000 (*M_r* 19K) was generated (Fig. 3), which is the same size as the predominant 19K PrP^{Sc} species generated in scrapie-infected neuroblastoma cells when PrP^{Sc} is labelled in the presence of tunicamycin¹⁶. This provided additional evidence that the cell-free conversion of PrP^C to PK-resistant forms mimicked that which occurs in scrapie-infected cells and did not require *N*-linked glycosylation of PrP^C.

To test the specificity of this conversion, this reaction was also done in the presence of a complex mixture of ³⁵S-labelled proteins (Fig. 4). Virtually all of these proteins were thoroughly digested by PK to small fragments running at the dye front of the gel regardless of the presence of PrP^{Sc}. Only one 17K

FIG. 2 Conversion of ³⁵S-PrP^C to PK-resistant forms in the presence of unlabelled PrP^{Sc}. *a*, SDS-PAGE-fluorography analysis of ³⁵S-PrP incubated with or without unlabelled PrP^{Sc} (lanes 1–10). The PrP^{Sc} was pretreated with the designated concentration of GdnHCl before incubation with ³⁵S-PrP^C. The ³⁵S-PrP^C was recombinant hamster PrP expressed in mouse neuroblastoma cells²⁹ which, without PK treatment, is composed of monomeric (30–40K) and apparently dimeric (60K) PrP bands on SDS-PAGE (lane 1) (S.A.P., B. Caughey and B.C., manuscript in preparation). Lanes 2–10, PK-treated samples representing 7-fold more ³⁵S-PrP^C equivalents than shown in lane 1. Lanes 11 and 12, immunoblot of PK-treated PrP^{Sc} species isolated from scrapie-infected mouse neuroblastoma cells (NB)²³ and hamster brain (Br)¹¹, respectively. The differences in the band pattern of these PrP^{Sc} samples are largely due to the variable glycosylation of PrP species from different sources. Lane 13, immunoblot of hamster brain PrP^{Sc} without PK treatment. *b*, Conversion of separate 30–40K (lanes 1 and 2) and 60K (lanes 3 and 4) ³⁵S-PrP^C species to PK resistant forms (lanes 5–8) after incubation with PrP^{Sc} pretreated with 3 M GdnHCl as in *a*. The 30–40K PrP^C was isolated by selectively releasing it from intact cells with phosphatidylinositol-specific phospholipase C before immunoprecipitation; the 60K PrP^C is resistant to this treatment (S.A.P., B. Caughey and B.C., manuscript in preparation). The 60K form aggregates and was selectively pelleted from cell lysates by ultracentrifugation before immunoprecipitation. *c*, Same as *a* except that a mutated hamster ³⁵S-PrP^C expressed in mouse fibroblasts was used which lacked the C-terminal signal sequence for the attachment of the glycoprophosphatidylinositol (GPI) membrane anchor. The predominant 24K species of this PrP without PK treatment (lane 1) was unglycosylated (G. Katzenstein and B. Caughey, unpublished observations). The major PK-resistant ³⁵S-PrP bands comigrating with the 17K-*M_r* marker is the same size as PrP^{Sc} generated in scrapie-infected NB cells expressing a similar GPI-less PrP construct³⁰ and close in migration to the 19K unglycosylated mouse PrP^{Sc} from scrapie-infected NB cells¹⁶ shown in lane 11. As much as a third of the GPI-less ³⁵S-PrP^C was converted to PK-resistant forms in some experiments (data not shown) and was routinely more efficiently converted than the complete, more heavily glycosylated PrP^C used in *a*. This is consistent with previous cell culture studies suggesting that less glycosylated PrP^C species preferentially converted to PK-resistant forms⁸. METHODS. The two types of PrP^C were metabolically labelled with ³⁵S-methionine and immunoprecipitated from detergent lysates of the



cells as described previously²³ except for the use of a monoclonal antibody (3F4) specific for the recombinant hamster PrP species. The ³⁵S-PrP^C was eluted from the immunoprecipitate in 3 M GdnHCl and mixed into the PrP^{Sc} renaturation reaction at an approximate PrP^C:PrP^{Sc} ratio of $\leq 1:50$ without altering the final concentrations of the other constituents described in Fig. 1 legend except that PrP^{Sc} pretreated with 6 M GdnHCl was also used. The 6 M GdnHCl was diluted to 1.1 M (instead of 0.75 M) for the incubation before PK treatment. In additional experiments, similar results were obtained when the mixture was diluted to 0.75 M GdnHCl (data not shown). The PK treatments and subsequent electrophoretic analyses were as described for Fig. 1 except that the gels were developed for the fluorographic detection of ³⁵S-labelled proteins²³. The PrP^C construct lacking the C-terminal GPI signal sequence was generated as follows. Codon 231 of the hamster PrP complementary DNA sequence was replaced with a stop codon. The mutated gene was then cloned into the pSFF retroviral expression vector by standard procedures and expressed in mixed cultures of $\Psi 2$ and PA317 packing cells derived from NIH 3T3 fibroblasts²⁹. A subclone of this mixed culture was used for the metabolic labelling of the recombinant PrP^C.

PK-resistant band was observed; this seemed to be PrP because it comigrated with the PK-resistant PrP band produced in the positive control reaction and was formed only in the 2-day reaction containing both ^{35}S -PrP^c and PrP^{Sc}. As another specificity control, cerebrovascular β -amyloid isolated from an Alzheimer's disease patient was substituted for PrP^{Sc}, at the same protein concentration, in the conversion reactions described in Fig. 2a and c, but no specific PK-resistant ^{35}S -labelled PrP species were observed (data not shown).

Our study shows that PrP^c can be converted selectively to PK-resistant forms similar to PrP^{Sc} in a cell-free system. The products of the conversion reaction were also much more resistant to digestion with other proteases, that is, trypsin and chymotrypsin, than was the PrP^c precursor, giving resistant fragments at least as large as those generated by PK (data not shown). The major protein constituents of the conversion reaction are ^{35}S -PrP^c, the monoclonal antibody used for immunoprecipitation and unlabelled PrP^{Sc}. The antibody is unlikely to be involved in PK-resistant PrP formation because incubation of ^{35}S -PrP^c with antibody alone did not induce the conversion to the large protease-resistant PrP species (Fig. 2). Therefore direct interactions between PrP^c and PrP^{Sc} are probably sufficient for the conversion to occur. This observation is consistent with certain 'protein only' models for PrP^{Sc} formation^{3,6,10,17-19}. However, our results do not rule out the involvement of other non-PrP constituents (for example, glycosaminoglycans and nucleic acids) that have been found associated with PrP^{Sc} (refs 20-22) and might be cofactors in the PrP^c to PrP^{Sc} conversion^{23,24}. Nor is it possible to exclude an *in vivo* role of an exogenous virus-like agent in producing conditions optimal for PrP^{Sc} formation, for example by upregulation of PrP^c or alteration of PrP^c catabolism as a secondary response to infection.

Nevertheless, the present study shows clearly that PrP^c from uninfected sources can be converted to PK-resistant forms by

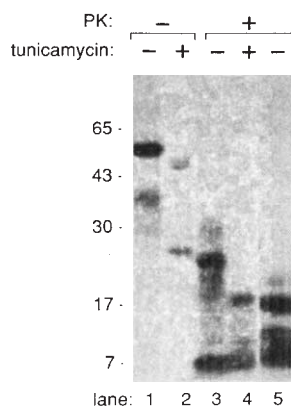


FIG. 3 Conversion of unglycosylated PrP^c to PK-resistant forms. Lanes 1-4, the full hamster PrP^c construct used in Fig. 2a was labelled with or without $15 \mu\text{g ml}^{-1}$ tunicamycin included in the preincubation and labelling media. The PrP^c was immunoprecipitated, mixed with PrP^{Sc} and analysed after a 2-day incubation with and without PK treatment as described for Fig. 2a. The PK-treated lanes contained 7-fold more ^{35}S -PrP^c equivalents than the non-PK-treated lanes. The prevention of N-linked glycosylation by tunicamycin yielded non-PK-treated species at 25K and ~50K which probably represent unglycosylated monomeric¹⁵ and dimeric forms of PrP, respectively (lane 2). On digestion of the tunicamycin-treated PrP with PK, a prominent band at ~19K is generated which is the same size as the lowest, unglycosylated PrP^{Sc} band shown in lane 11 of Fig. 2a, c and migrated slightly ahead of the major 17K PK-resistant band (lane 5) derived from the conversion of the GPI-less, mostly unglycosylated PrP^c described for Fig. 2c. These results provided evidence that the larger (>19K) ^{35}S -labelled PK-resistant PrP species generated without the tunicamycin treatment (lane 3) contained N-linked glycans and a protein core encompassing the N-linked glycosylation sites at residues 181 and 197.

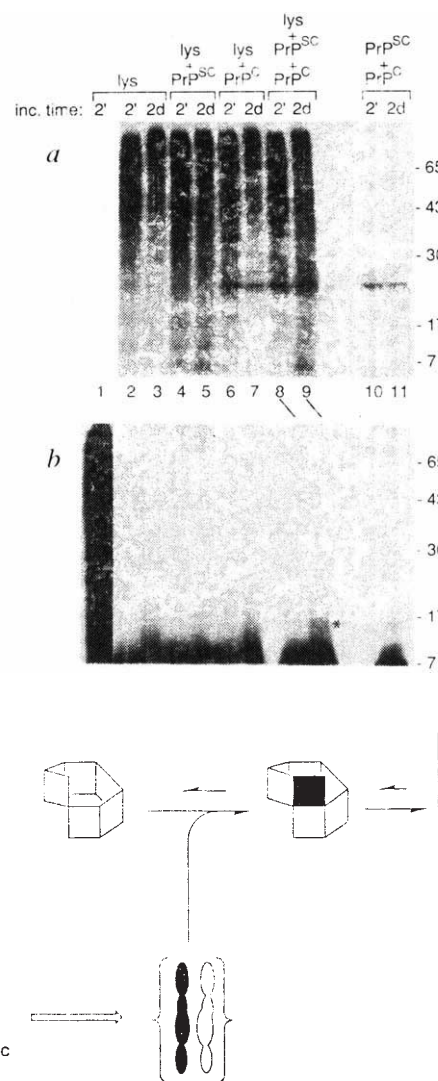


FIG. 4 Specificity of the conversion of PrP^c to PK-resistant forms in the presence of PrP^{Sc}. ^{35}S -labelled fibroblast cell lysate proteins (minus hamster PrP^c) ('lys') were tested for the conversion to PK-resistant forms. a, SDS-PAGE-fluorography of combinations of ^{35}S -labelled lysate proteins, PrP^{Sc} and ^{35}S -PrP^c (lacking the GPI signal as described for Fig. 2c) incubated for the designated times (inc. time). When ^{35}S -PrP^c was added, the ratio of ^{35}S -labelled lysate protein to ^{35}S -PrP^c radioactivity (c.p.m.) was 15:1. b, Same as a except that the samples were also treated with PK (except for lane 1). Equivalent proportions of the reaction mixtures are represented in a and b. However, the fluorographic exposure times for a (all lanes), b (lanes 1-9) and b (lanes 10 and 11) were 2 h, 12 h and 48 h, respectively, in order to best visualize the bands of interest. Lane b1 is the same as a2 except for the 6 times longer exposure time to show the direct comparison of the control and PK-digested samples. The asterisk marks the position of the major 17K PK-resistant ^{35}S -PrP band generated in lanes 9 and 11.

METHODS. The protocol and conditions were described in Figs 1 and 2 legends except for the addition of the fibroblast cell lysate proteins remaining in the supernatant of the hamster PrP^c immunoprecipitates. A methanol precipitate of the lysate proteins was dissolved in 3 M GdnHCl and mixed with ^{35}S -PrP^c and PrP^{Sc} in 3 M GdnHCl before the first dilution and incubation steps. c, A working hypothesis to explain the selective cell-free conversion of ^{35}S -PrP^c (black circle) to PrP^{Sc} (squares). Denatured (or partially denatured) GdnHCl-treated ^{35}S -PrP^c is added to partially denatured or native PrP^{Sc} under conditions where denaturation is reversible. The ^{35}S -PrP^c interacts specifically with the PrP^{Sc} seed and is altered conformationally and incorporated into the protease-resistant aggregate^{19,25}. Under conditions where PrP^{Sc} is completely denatured, very little incorporation is observed, suggesting that *in situ* seeding is required.

means that do not require the biosynthesis of new macromolecules. These results and studies of amyloid fibril formation by peptide models of PrP (ref. 25) are consistent with PrP^{Sc} being made by a seeded polymerization mechanism similar to that of amyloid fibril formation by other proteins (Figs 1b and 4c)¹⁹. Thus, it is possible that there are similar underlying mechanisms for the pathogenic accumulation of proteins associated with TSEs and other amyloid diseases¹⁹. Nonetheless, there are important differences between these disease types because amyloidoses are not generally viewed as transmissible and TSEs do not always involve amyloid formation by PrP^{Sc} (ref. 26). The present experimental model of PrP^{Sc} formation should greatly facilitate detailed analyses of the essential factors and mechanisms in this central event in TSE pathogenesis. One limitation of the current conversion reaction is that the PrP^{Sc} 'seed' is in vast (≥ 50 -fold) excess over PrP^C. Thus, the preexisting scrapie infectivity in the present system would overwhelm the detection of any new scrapie infectivity associated with the conversion of tracer levels of ³⁵S-PrP^C to PK-resistant forms. However, further refinements of this cell-free system may finally allow a direct assessment of the relationship of this PrP conversion to scrapie infectivity. □

5. Düringer, H. et al. *Nature* **306**, 476–478 (1983).
6. Prusiner, S. B. *Science* **252**, 1515–1522 (1991).
7. Meyer, R. K. et al. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2310–2314 (1986).
8. Caughey, B. & Raymond, G. J. *J. biol. Chem.* **266**, 18217–18223 (1991).
9. Borchelt, D. R., Scott, M., Taraboulos, A., Stahl, N. & Prusiner, S. B. *J. Cell Biol.* **110**, 743–752 (1990).
10. Hope, J. et al. *EMBO J.* **5**, 2591–2597 (1986).
11. Caughey, B. W. et al. *Biochemistry* **30**, 7672–7680 (1991).
12. Pan, K.-M. et al. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10962–10966 (1993).
13. Stahl, N. et al. *Biochemistry* **32**, 1991–2002 (1993).
14. Safar, J., Roller, P. P., Gajdusek, D. C. & Gibbs, C. J. *J. biol. Chem.* **268**, 20276–20284 (1993).
15. Caughey, B., Race, R. E., Ernst, D., Buchmeier, M. J. & Chesebro, B. J. *Virology* **63**, 175–181 (1989).
16. Taraboulos, A. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 8262–8266 (1990).
17. Bolton, D. C. & Benheim, P. E. In *Novel Infectious Agents and the Central Nervous System* Vol. 1 (eds Bock, G. & Marsh, J.) 164–181 (Wiley, Chichester, 1988).
18. Brown, P., Goldfarb, L. G. & Gajdusek, D. C. *Lancet* **337**, 1019–1022 (1991).
19. Jarrett, J. T. & Lansbury, P. T. Jr *Cell* **73**, 1055–1058 (1993).
20. Snow, A. D. et al. *Lab. Invest.* **63**, 601–611 (1990).
21. Aiken, J. M. & Marsh, R. F. *Microbiol. Rev.* **54**, 242–246 (1990).
22. Riesner, D. *Semin. Virol.* **2**, 215–226 (1991).
23. Caughey, B. & Raymond, G. J. *J. Virol.* **67**, 643–650 (1993).
24. Weissmann, C. *Nature* **352**, 679–683 (1991).
25. Come, J. H., Fraser, P. E. & Lansbury, P. T. Jr *Proc. natn. Acad. Sci. U.S.A.* **90**, 5959–5963 (1993).
26. Fraser, H. *Br. med. Bull.* **49**, 792–809 (1993).
27. Prusiner, S. B., Groth, D. F., Bolton, D. C., Kent, S. B. & Hood, L. E. *Cell* **38**, 127–134 (1984).
28. Caughey, B., Raymond, G. J., Ernst, D. & Race, R. E. *J. Virol.* **65**, 6597–6603 (1991).
29. Chesebro, B. et al. *Dev. biol. Stand.* **80**, 131–140 (1993).
30. Rogers, M., Yehli, F., Scott, M. & Prusiner, S. B. *Proc. natn. Acad. Sci. U.S.A.* **90**, 3182–3186 (1993).

Received 3 May; accepted 30 June 1994.

1. Rohwer, R. G. *Curr. Topics Microbiol. Immun.* **172**, 195–232 (1991).
2. Pattison, I. H. & Jones, K. M. *Vet. Rec.* **80**, 2–9 (1967).
3. Griffith, J. S. *Nature* **215**, 1043–1044 (1967).
4. Bolton, D. C., McKinley, M. P. & Prusiner, S. B. *Science* **218**, 1309–1311 (1982).

ACKNOWLEDGEMENTS. We thank G. Glenner for providing the Alzheimer's cerebrovascular β -amyloid, G. Hettrick and B. Evans for graphics assistance and G. Katzenstein for tissue culture assistance. P.T.L. is the recipient of a Presidential Young Investigator Award from the National Science Foundation.

The inositol trisphosphate calcium channel is inactivated by inositol trisphosphate

György Hajnóczky & Andrew P. Thomas*

Department of Pathology and Cell Biology,
Thomas Jefferson University, Philadelphia, Pennsylvania 19107, USA

ACTIVATION of intracellular Ca^{2+} channels by inositol 1,4,5-trisphosphate ($\text{Ins}(1,4,5)\text{P}_3$) represents the initial Ca^{2+} mobilization step in response to many extracellular signals¹. Here we report that $\text{Ins}(1,4,5)\text{P}_3$ -induced channel activation in permeabilized hepatocytes is followed by a time-dependent inactivation, which is a direct consequence of ligand binding. The inactivation by $\text{Ins}(1,4,5)\text{P}_3$ parallels the quantal character of channel opening, giving rise to a unique process of incremental inactivation whereby discrete channel populations are inhibited at each $\text{Ins}(1,4,5)\text{P}_3$ dose. $\text{Ins}(1,4,5)\text{P}_3$ can induce inactivation in the absence of stored Ca^{2+} , but the inactivation rate is enhanced by increases of cytosolic Ca^{2+} . The inhibitory effect of $\text{Ins}(1,4,5)\text{P}_3$ can be reversed by $\text{Ins}(1,4,5)\text{P}_3$ washout, or by chelation of cytosolic Ca^{2+} . Thus, $\text{Ins}(1,4,5)\text{P}_3$ and Ca^{2+} act as coinhibitors of the $\text{Ins}(1,4,5)\text{P}_3$ -sensitive Ca^{2+} channel. Inactivation is an inherent consequence of $\text{Ins}(1,4,5)\text{P}_3$ -induced channel opening which can terminate increases of cytosolic Ca^{2+} .

The permeability of $\text{Ins}(1,4,5)\text{P}_3$ -sensitive Ca^{2+} channels was monitored in permeabilized hepatocytes by measuring $\text{Ins}(1,4,5)\text{P}_3$ -induced entry of tracer Mn^{2+} into intracellular stores loaded with Fura2, where it was detected by the essentially stoichiometric quench of fluorescence. The stores were depleted of Ca^{2+} by preincubation with thapsigargin. This method allowed net channel conductance to be followed throughout extended periods of $\text{Ins}(1,4,5)\text{P}_3$ exposure^{2–4}. Addition of Mn^{2+}

resulted in a prompt quench of released cytosolic Fura2, followed by a slow basal quench of compartmentalized dye (Fig. 1a). About 50% of the compartmentalized Fura2 was quenched rapidly when the $\text{Ins}(1,4,5)\text{P}_3$ -sensitive channels were activated with maximal $\text{Ins}(1,4,5)\text{P}_3$. To examine the permeability properties of the channel after various periods of $\text{Ins}(1,4,5)\text{P}_3$ exposure, $\text{Ins}(1,4,5)\text{P}_3$ and Mn^{2+} were added simultaneously to permeabilized cells preincubated in the absence or presence of $\text{Ins}(1,4,5)\text{P}_3$ (Fig. 1a). Addition of $\text{Ins}(1,4,5)\text{P}_3$ before Mn^{2+} did not evoke any fluorescence change², but markedly inhibited the subsequent rate of $\text{Ins}(1,4,5)\text{P}_3$ -dependent Mn^{2+} quench. The time course of channel inactivation under these conditions was reduced 10-fold within 60 s, and this was half-maximal at 15 s (Fig. 1b). This time-dependent decline in channel activity was not secondary to Ca^{2+} fluxes, because these were eliminated by thapsigargin^{2,3}.

$\text{Ins}(1,4,5)\text{P}_3$ -induced Mn^{2+} quench in naive cells did not follow a single exponential, but gave an excellent biexponential fit (Fig. 1c) with rate constants $0.55 \pm 0.13 \text{ s}^{-1}$ and $0.06 \pm 0.02 \text{ s}^{-1}$ and pool sizes $10.6 \pm 0.6\%$ and $9.2 \pm 0.2\%$ ($n=4$). These data are consistent with measurements of $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} efflux in several cell types that have revealed two kinetic components differing by 10–30-fold in rate constant^{5,6}. As noted above, feedback effects resulting from Ca^{2+} release can be excluded here. Thus, the two rate constants may reflect distinct $\text{Ins}(1,4,5)\text{P}_3$ receptor subtypes, or the slow component could represent an inactivated form of the same channel. The rapid component was absent after $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation, which resulted in monoexponential kinetics of Mn^{2+} quench with rate constant $0.02 \pm 0.01 \text{ s}^{-1}$ and pool size $16.4 \pm 1.4\%$. These data provide further evidence for a large shift in the net conductance of $\text{Ins}(1,4,5)\text{P}_3$ receptor channels in the sustained presence of $\text{Ins}(1,4,5)\text{P}_3$.

Preincubation with submaximal $\text{Ins}(2,4,5)\text{P}_3$ (used to minimize metabolism) reduced the initial rate of Mn^{2+} quench compared with the rate with $\text{Ins}(2,4,5)\text{P}_3$ and Mn^{2+} added simultaneously, but did not cause similar inactivation to maximal $\text{Ins}(1,4,5)\text{P}_3$ (Fig. 2a). The Mn^{2+} quench rate reflects the

* To whom correspondence should be addressed.

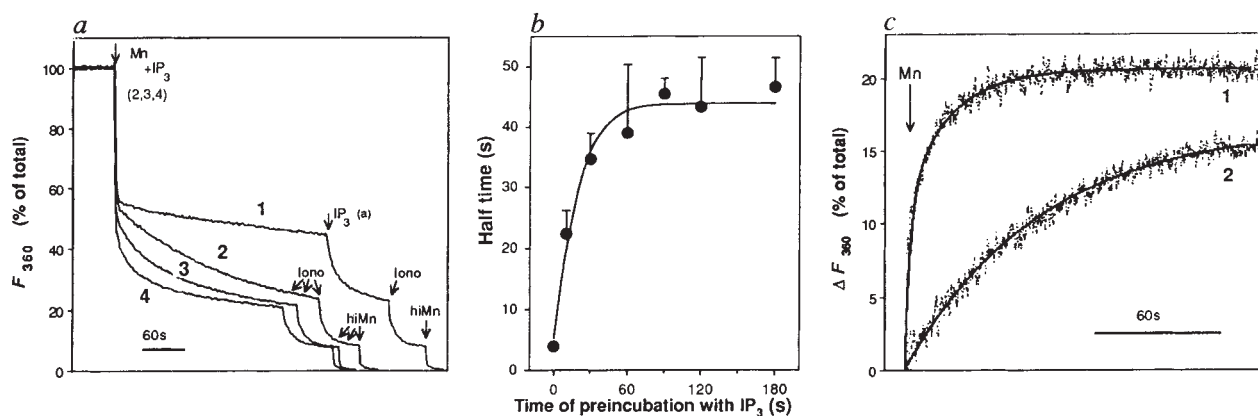


FIG. 1 Decline of Ins(1,4,5)P₃-dependent Mn²⁺ permeability in the sustained presence of Ins(1,4,5)P₃. **a**, Fluorescence quenching was initiated by addition of 40 μM MnCl₂ (2 μM free Mn²⁺, buffered by ATP) 7 min after permeabilization and the effects of 7.5 μM Ins(1,4,5)P₃ were examined as follows: trace 1, Mn²⁺ quench in the absence of Ins(1,4,5)P₃ followed by Ins(1,4,5)P₃ addition 330 s later; traces 2–4, Ins(1,4,5)P₃ added simultaneously with Mn²⁺ after Ins(1,4,5)P₃ preincubation for 180 s (2), 20 s (3) or no Ins(1,4,5)P₃ prepulse (4). At the end of each run, 4 μM ionomycin (iono) and 500 μM MnCl₂ (hiMn) were added to quench residual Fura2 in InsP₃-insensitive compartments. Data are expressed as a percentage of initial Mn²⁺-sensitive fluorescence. **b**, Time course of Ins(1,4,5)P₃-induced inactivation determined using the protocol shown in **a**. The $t_{1/2}$ of Ins(1,4,5)P₃-dependent quench (measured after subtraction of the quench without Ins(1,4,5)P₃) is plotted as a function of the duration of Ins(1,4,5)P₃ prepulse (mean ± s.e., $n=4$). **c**, The Ins(1,4,5)P₃-dependent component of Mn²⁺ quench for cells preincubated in the absence (trace 1) or presence of 7.5 μM Ins(1,4,5)P₃ for 180 s (trace 2) was obtained by subtraction of a parallel quench response in the absence of Ins(1,4,5)P₃. Solid lines

show nonlinear regression fits.

METHODS. Freshly isolated hepatocytes were loaded with Fura2/AM using a protocol that resulted in ~50% compartmentation of Fura2 in intracellular stores^{2,3}. Washed cells were resuspended in medium composed of 120 mM KCl, 20 mM HEPES/Tris, 10 mM NaCl, 1 mM KH₂PO₄, 2 mM Mg-ATP, 5 mM phosphocreatine, 1 U ml⁻¹ creatine kinase, 1 μg ml⁻¹ each of antipain, leupeptin, pepstatin, 1 μM ruthenium red and 2 μM thapsigargin, pH 7.2, and immediately permeabilized with 25 μg ml⁻¹ digitonin. The cells were incubated at 35 °C in a rapidly stirred fluorometer cuvette^{2,3}. Fura2 fluorescence was excited at 360 nm (Ca²⁺-insensitive wavelength) and detected at 510 nm. Medium [Ca²⁺] was 300 nM (estimated with the released Fura2). In the presence of thapsigargin the [Ca²⁺] in the intracellular stores was depleted close to the concentration in the medium, as indicated by the small decrement in 340 nm Fura2 fluorescence following Ins(1,4,5)P₃ addition to cells in medium with Ca²⁺ buffered by 1 mM EGTA. Metabolism of Ins(1,4,5)P₃ was slow under the conditions used for these experiments (20% decrease of 125 nM [³H]Ins(1,4,5)P₃ in 3 min) and the inactivation was mimicked by Ins(2,4,5)P₃ but not by Ins(1,3,4,5)P₄.

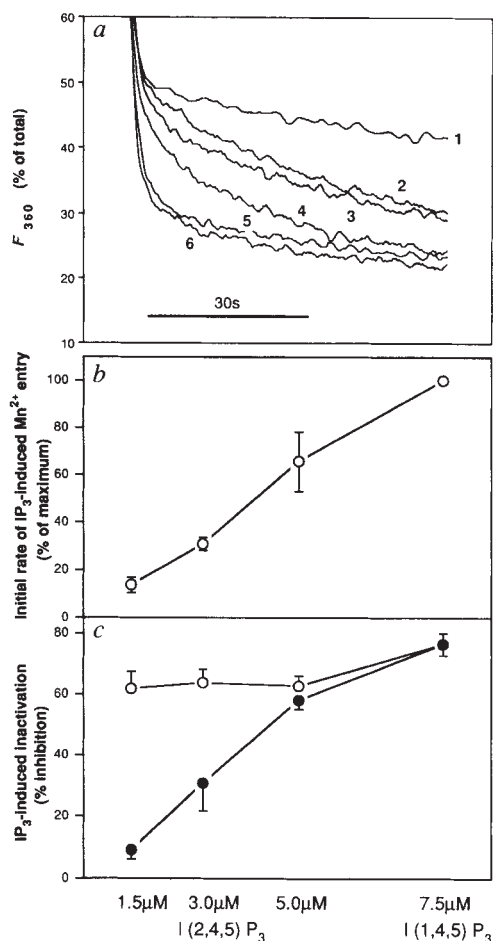


FIG. 2 Incremental inactivation by submaximal Ins(1,4,5)P₃. **a**, Fluorescence traces show quench response following addition of 100 μM MnCl₂: trace 1, no other additions; 2, 1.5 μM Ins(2,4,5)P₃ 90 s before Mn²⁺; 3, 1.5 μM Ins(2,4,5)P₃ added with Mn²⁺; 4, 7.5 μM Ins(1,4,5)P₃ 90 s before Mn²⁺; 5, 1.5 μM Ins(2,4,5)P₃ for 90 s then 7.5 μM Ins(1,4,5)P₃ added with Mn²⁺; 6, 7.5 μM Ins(1,4,5)P₃ added with Mn²⁺. Note that 1.5 μM Ins(2,4,5)P₃, which gave about 15% of maximal quench rate, effectively inactivated itself (traces 2 versus 3) but only slightly reduced the response to maximal Ins(1,4,5)P₃ (traces 5 versus 6). **b**, Dose-dependence of quench rate initiated by addition of inositol phosphates to naive cells in the presence of Mn²⁺. Values are corrected for the quench rate in unstimulated cells. **c**, Experiments were as shown in **a** to measure the effect of activating fractions of the Ins(1,4,5)P₃-sensitive channels during the 90-s prepulse. The initial quench rate was measured either without further inositol phosphate addition (○) or with maximal Ins(1,4,5)P₃ (7.5 μM) added together with the Mn²⁺ (●). The data are expressed as percentage inhibition from the control rate measured with the relevant inositol phosphates added simultaneously with Mn²⁺. Values are mean ± s.e. with $n=3-6$. Ins(2,4,5)P₃, which has a lower affinity for the receptor than Ins(1,4,5)P₃, was used to minimize metabolism at the submaximal doses, but similar results were obtained with Ins(1,4,5)P₃. The 5-phosphatase inhibitor (1R, 2R, 4R)-cyclohexane-1,2,4-tris-(methylene-sulphonate) (16 μM) was also included in all these experiments.

fractional activation of $\text{Ins}(1,4,5)\text{P}_3$ receptors²⁻⁴, which increased with $\text{Ins}(2,4,5)\text{P}_3$ dose (Fig. 2b). When constant $\text{Ins}(2,4,5)\text{P}_3$ doses were used during both the prepulse and subsequent Mn^{2+} quench the percentage inactivation was similar at each dose, whereas the response to maximal $\text{Ins}(1,4,5)\text{P}_3$ was inhibited in proportion to the fraction of channels activated during the prepulse period (Fig. 2c). Thus, inactivation was limited to only the occupied fraction of receptors and high $\text{Ins}(1,4,5)\text{P}_3$ did not overcome the inhibition of this channel population. These data indicate that $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation occurs in an incremental manner, which parallels the quantal, or incremental, nature of $\text{Ins}(1,4,5)\text{P}_3$ -induced channel opening^{2,3,6-11}. Incremental Ca^{2+} release can be eliminated at low temperature^{6,11} and, consistent with this, we found that the inactivation by $\text{Ins}(1,4,5)\text{P}_3$ also does not occur at 4 °C (not shown).

$\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} release is affected by cytosolic Ca^{2+} in a biphasic manner, with first potentiation and then inactivation at higher Ca^{2+} concentrations¹²⁻¹⁶. We therefore examined the effect of varying cytosolic Ca^{2+} during the prepulse period

on $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation. Concentrations of 350 nM Ca^{2+} and below did not inactivate in the absence of $\text{Ins}(1,4,5)\text{P}_3$ (Fig. 3a, traces labelled 3). However, inactivation by $\text{Ins}(1,4,5)\text{P}_3$ was dependent on the presence of low levels of Ca^{2+} (Fig. 3a, traces labelled 2). When cells were prepulsed for 20 s with maximal $\text{Ins}(1,4,5)\text{P}_3$, the $\text{Ins}(1,4,5)\text{P}_3$ -dependent inactivation increased progressively over the physiological range of $[\text{Ca}^{2+}]_i$ from 30 nM to 1 μM (Fig. 3b). Although Ca^{2+} did not inhibit independent of $\text{Ins}(1,4,5)\text{P}_3$ at ≤ 350 nM, preincubation with 1 μM Ca^{2+} resulted in partial inhibition (19%) of the subsequent $\text{Ins}(1,4,5)\text{P}_3$ -induced Mn^{2+} quench rate. Nevertheless, inclusion of $\text{Ins}(1,4,5)\text{P}_3$ in the prepulse caused additional inhibition even at the highest Ca^{2+} levels examined (Fig. 3b). Thus, the effect of Ca^{2+} alone could actually reflect an $\text{Ins}(1,4,5)\text{P}_3$ -dependent inhibition which develops immediately after $\text{Ins}(1,4,5)\text{P}_3$ is added. Alternatively, $\text{Ins}(1,4,5)\text{P}_3$ may induce channel inactivation by aiding access of Ca^{2+} to its inhibitory site.

$\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation was reversible. When cells were prepulsed with $\text{Ins}(1,4,5)\text{P}_3$ for 60 s followed by ligand

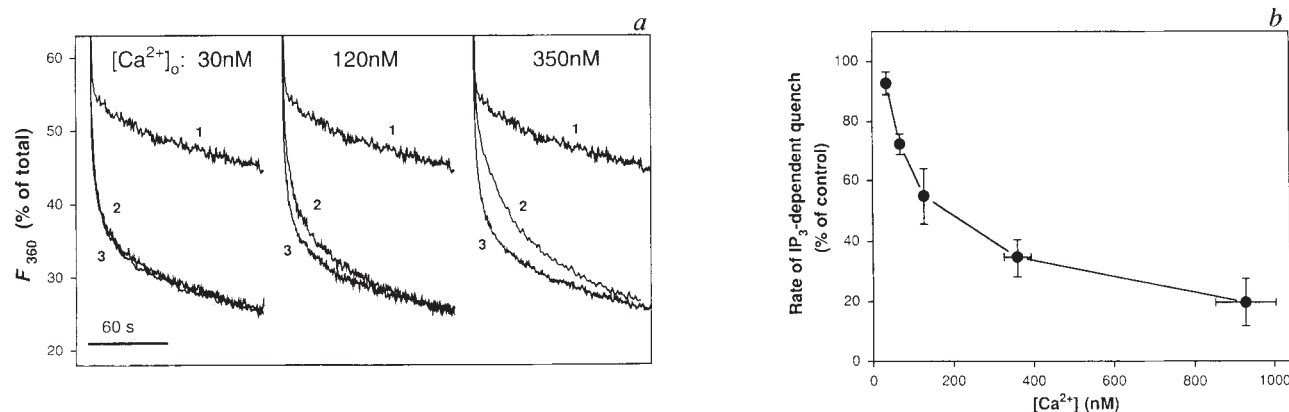


FIG. 3 Ca^{2+} -dependence of $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation. The effect of a 20-s prepulse with 7.5 μM $\text{Ins}(1,4,5)\text{P}_3$ was examined as described for Fig. 1, except that $[\text{Ca}^{2+}]$ in the preincubation medium was varied by addition of EGTA (5–23 μM) and CaCl_2 (8 μM). Free Ca^{2+} was measured using the released cytosolic Fura2. Mn^{2+} quench was initiated with 60 μM MnCl_2 added together with EGTA to give the same concentration during the quench phase of each run. a, Mn^{2+} quench

after preincubation at 30, 120 and 350 nM Ca^{2+} in the absence (traces 3) and presence (traces 2) of $\text{Ins}(1,4,5)\text{P}_3$. Traces (1) show $\text{Ins}(1,4,5)\text{P}_3$ -independent quench controls. b, The $\text{Ins}(1,4,5)\text{P}_3$ -dependent quench rate measured over the first 3 s was expressed as a percentage of the $\text{Ins}(1,4,5)\text{P}_3$ -induced quench rate for cells preincubated at the same Ca^{2+} level without $\text{Ins}(1,4,5)\text{P}_3$. Values are mean $\pm$ s.e. of data from at least 4 experiments.

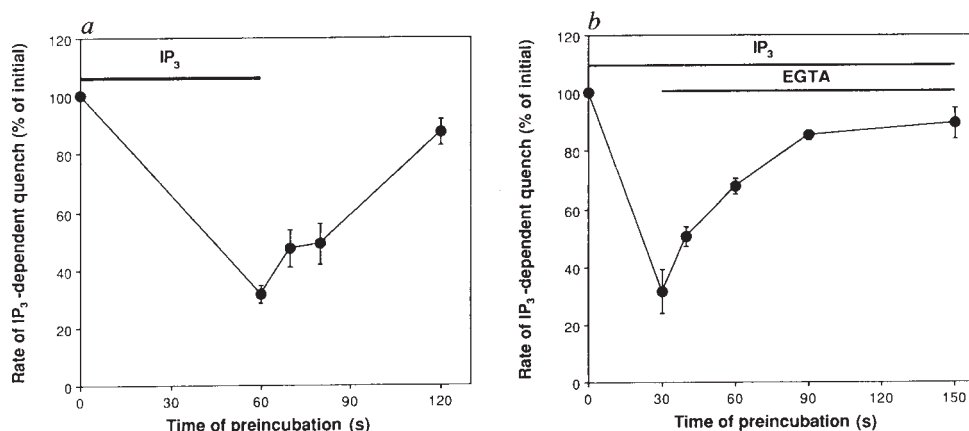


FIG. 4 Reversibility of $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation. a, Permeabilized cells were preincubated for 60 s in the presence or absence of 0.5 μM $\text{Ins}(1,4,5)\text{P}_3$. The $\text{Ins}(1,4,5)\text{P}_3$ concentration was then reduced by 5-fold dilution and $\text{Ins}(1,4,5)\text{P}_3$ -induced Mn^{2+} quench was measured at the indicated times by addition of 7.5 μM $\text{Ins}(1,4,5)\text{P}_3$ and 40 μM MnCl_2 . The initial quench rate over the first 3 s is expressed as a percentage of the $\text{Ins}(1,4,5)\text{P}_3$ -dependent quench rates for cells preincuba-

ted without $\text{Ins}(1,4,5)\text{P}_3$. The extent of $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation was the same in concentrated and dilute cell suspensions. b, Permeabilized cells were preincubated with 7.5 μM $\text{Ins}(1,4,5)\text{P}_3$ for 30 s, followed by addition of 100 μM EGTA. Quenching was initiated by 150 μM MnCl_2 in the presence of 100 μM EGTA and 7.5 μM $\text{Ins}(1,4,5)\text{P}_3$. Values are mean $\pm$ s.e. with $n=3$ in each case.

dilution and a further period of preincubation, the $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation recovered over a period of about 60 s (Fig. 4a). A time-dependent recovery from inactivation could also be induced by chelation of medium Ca^{2+} with EGTA in the sustained presence of saturating $\text{Ins}(1,4,5)\text{P}_3$ (Fig. 4b). This again shows the interdependent nature of inhibition by $\text{Ins}(1,4,5)\text{P}_3$ and Ca^{2+} and demonstrates that recovery from the inactivated state can occur as a result of decreased levels of either $\text{Ins}(1,4,5)\text{P}_3$ or Ca^{2+} .

The present study represents the first demonstration that the $\text{Ins}(1,4,5)\text{P}_3$ receptor Ca^{2+} channel is inactivated as a result of $\text{Ins}(1,4,5)\text{P}_3$ binding. This observation has been helped by the Mn^{2+} quench technique, which permits dissociation of the permeability measurement from $\text{Ins}(1,4,5)\text{P}_3$ binding and Ca^{2+} flux events, and allows ion permeability to be studied *in situ* over an extended period. Some earlier reports have concluded that $\text{Ins}(1,4,5)\text{P}_3$ does not desensitize its receptor^{6,10,12}. However, in these studies the measurements of channel activity relied on $\text{Ins}(1,4,5)\text{P}_3$ -induced Ca^{2+} fluxes that are short lived and known to have feedback effects on the receptor. Moreover, the data presented here demonstrate that $\text{Ins}(1,4,5)\text{P}_3$ -induced inactivation occurs in a novel incremental manner, whereby submaximal $\text{Ins}(1,4,5)\text{P}_3$ doses inactivate only the subpopulation of channels responding to that $\text{Ins}(1,4,5)\text{P}_3$ dose, rather than causing global inactivation of all receptors. In the intact cell, $\text{Ins}(1,4,5)\text{P}_3$ -

induced inactivation may provide a mechanism for adaptation to partially elevated basal levels of $\text{Ins}(1,4,5)\text{P}_3$. However, inactivation is likely to be manifest most strongly once channel opening is initiated and the process is potentiated by the interaction between $\text{Ins}(1,4,5)\text{P}_3$ and released Ca^{2+} . This could serve to limit the duration of channel opening and may play an important role in the termination phase of $[\text{Ca}^{2+}]$ spikes. □

Received 4 February; accepted 20 May 1994.

- Berridge, M. J. *Nature* **361**, 315–325 (1993).
- Hajnóczky, G., Gao, E., Nomura, T., Hoek, J. B. & Thomas, A. P. *Biochem. J.* **293**, 413–422 (1993).
- Renard-Rooney, D. C., Hajnóczky, G., Seitz, M. B., Schneider, T. G. & Thomas, A. P. *J. biol. Chem.* **268**, 23601–23610 (1993).
- Hajnóczky, G., Lin, C. & Thomas, A. P. *J. biol. Chem.* **269**, 10280–10287 (1994).
- Champeil, P. et al. *J. biol. Chem.* **264**, 17665–17673 (1989).
- Meyer, T. & Stryer, L. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3841–3845 (1990).
- Muallem, S., Pandol, S. J. & Beeker, T. G. *J. biol. Chem.* **264**, 205–212 (1989).
- Parker, I. & Ivorra, I. *Science* **250**, 977–979 (1990).
- Oldershaw, K. A., Nunn, D. L. & Taylor, C. W. *Biochem. J.* **278**, 705–708 (1991).
- Ferris, C. D., Cameron, A. M., Haganir, R. L. & Snyder, S. H. *Nature* **356**, 350–352 (1992).
- Kindman, L. A. & Meyer, T. *Biochemistry* **32**, 1270–1277 (1993).
- Finch, E. A., Turner, T. J. & Goldin, S. M. *Science* **252**, 443–446 (1991).
- Iino, M. *J. gen. Physiol.* **95**, 1103–1122 (1990).
- Bezprozvanny, I., Watras, J. & Ehrlich, B. E. *Nature* **351**, 751–754 (1991).
- Iino, M. & Endo, M. *Nature* **360**, 76–78 (1992).
- Marshall, I. C. B. & Taylor, C. W. *J. biol. Chem.* **268**, 13214–13220 (1993).

ACKNOWLEDGEMENTS. We thank B. Ehrlich, S. Joseph and J. Hoek for helpful suggestions. This work was supported by grants from the NIH.

Nucleosome disruption and enhancement of activator binding by a human SW1/SNF complex

Hyockman Kwon*, Anthony N. Imbalzano†, Paul A. Khavari‡, Robert E. Kingston† & Michael R. Green*§

* Howard Hughes Medical Institute, Program in Molecular Medicine, University of Massachusetts Medical Center, Worcester, Massachusetts 01605, USA

† Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

‡ Howard Hughes Medical Institute, Center for Molecular and Genetic Medicine, Stanford University, Stanford, California 94305, USA

CHROMATIN structure can affect the transcriptional activity of eukaryotic structural genes by blocking access of sequence-specific activator proteins (activators) to their promoter-binding sites¹. For example, the DNA-binding domain of the yeast GAL4 protein interacts very poorly with nucleosome cores compared with naked DNA² (and see below), and binding of other activators is even more strongly inhibited^{2,3}. The way in which activators bind to nucleosomal DNA is therefore a critical aspect of transcriptional activation. Genetic studies have suggested that the multi-component SWI/SNF complex of *Saccharomyces cerevisiae* facilitates transcription by altering the structure of the chromatin^{4,5}. Here we identify and partially purify a human homologue of the yeast SWI/SNF complex (hSWI/SNF complex). We show that a partially purified hSWI/SNF complex mediates the ATP-dependent disruption of a nucleosome, thereby enabling the activators, GAL4-VP16 and GAL4-AH, to bind within a nucleosome core. We conclude that the hSWI/SNF complex acts directly to reorganize chromatin structure so as to facilitate binding of transcription factors.

TABLE 1 Amount of GAL4 derivative required for half-maximal binding to nucleosomal templates in the presence of the hSWI/SNF complex

	Footprinting (U)		
	Expt. 1	Expt. 2	Expt. 3
GAL4-VP16	20	70	40
GAL4-AH	200	300	100
GAL4(1–94)	1,000	2,000	1,000

DNase I footprinting was as described for Fig. 2 using increasing amounts of each GAL4 derivative, 0.3 ng (1.12×10^{-10} M) specific DNA and 3 ng carrier DNA (*Hae*III-digested pUC18). Percentage protection was determined at each concentration of a GAL4 derivative by phosphorimager analysis as previously described¹⁹. The amount of each GAL4 derivative required for 50% protection of naked DNA was defined as 1 unit (U). Increasing amounts of each GAL4 derivative were then used to determine the amount required for 50% protection of nucleosomal DNA in the presence of 1.3 µg hSWI/SNF complex and 4 mM ATP. 50% protection of nucleosomal DNA was not achieved for any GAL4 derivative in the absence of ATP. *E. coli*-expressed GAL4 derivatives were purified as described^{20,21}.

Previous studies have shown that human cells contain homologues of the yeast protein, SWI2 (refs 6, 7). We therefore sought to identify and purify a SWI/SNF complex from a human (HeLa) cell nuclear extract. Our experimental strategy was to fractionate a nuclear extract and monitor the chromatographic fractions for the human SWI2 homologue. Figure 1a shows the chromatographic scheme and Fig. 1b the polypeptide composition of the fractions. The hSWI2 in a HeLa cell nuclear extract was present in two distinct phosphocellulose fractions (0.5 M and 0.7 M KCl) and is a component of two chromatographically separable hSWI2-containing complexes, designated hSWI/SNF complexes A and B.

Figure 1b shows the polypeptide composition of the two hSWI/SNF complexes after the final Superose 6 gel filtration (complex A, lane 1) or glycerol gradient (complex B, lane 2) step. In addition to hSWI2, two polypeptides of high relative molecular mass (M_r), p170 and p150, are present in both complexes. The hSWI/SNF complex A preparation contains four smaller polypeptides (p60, p50, p45 and p43) which were also detected in the purified hSWI/SNF complex B but seem to be present in lower relative amounts. These seven polypeptides

§ To whom correspondence should be addressed.

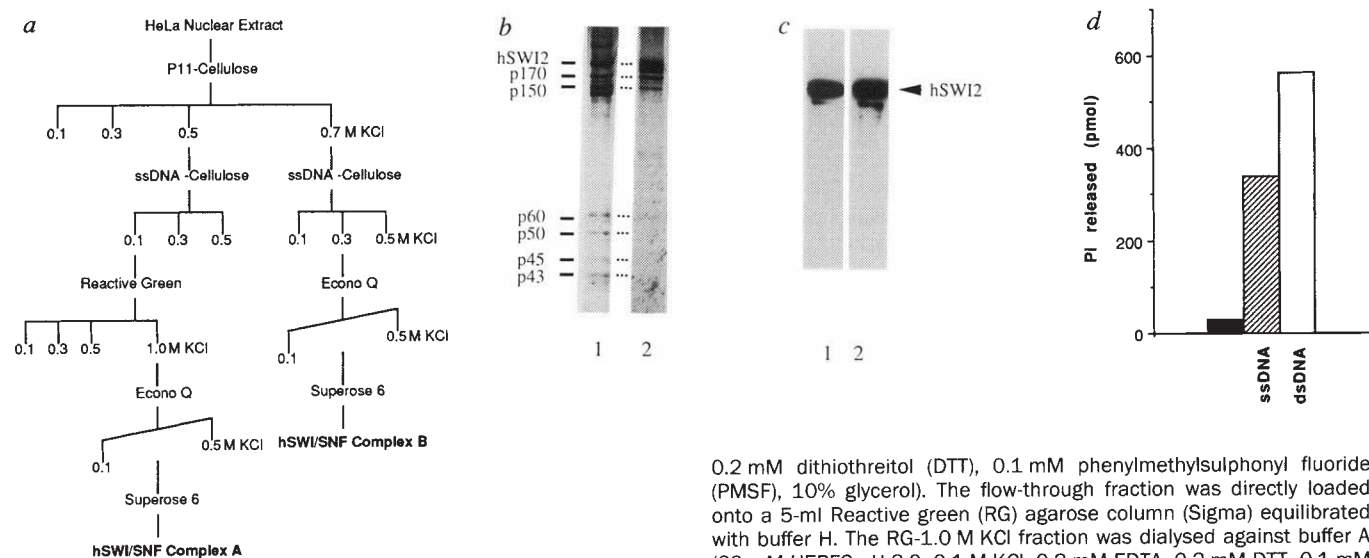
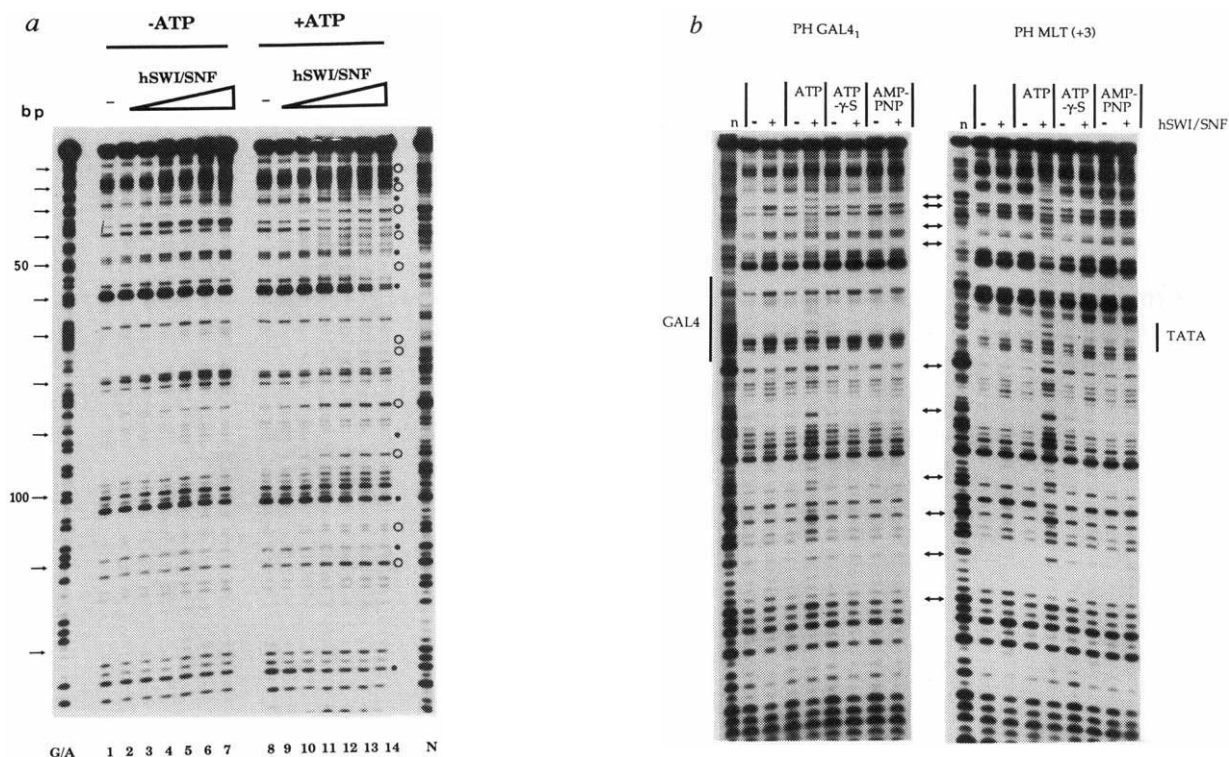


FIG. 1 Purification of hSWI/SNF complexes. **a**, Purification scheme; **b**, **c**, polypeptide composition of the chromatographic fractions. Silver staining (**b**) and an immunoblotting assay using the α -BRG1 antibody⁷ (**c**) of an SDS-PAGE gel showing the hSWI/SNF complex A (Superose 6 fraction) from the P11–0.5 M (lane 1) and the hSWI/SNF complex B (glycerol gradient fraction) from the P11–0.7 M KCl fractions (lane 2). **d**, The hSWI/SNF complex contains a DNA-dependent ATPase activity. ATP hydrolysis was measured in the absence (solid bar) or presence of 1 μ g single-stranded DNA (hatched bar) or double-stranded DNA (open bar). Stimulation by ssDNA and dsDNA was 11-fold and 19-fold, respectively. METHODS. The P11–0.5 M KCl fraction of HeLa cell nuclear extract (1 g) was applied to a ssDNA cellulose column (10 ml) equilibrated with buffer H (20 mM HEPES pH 8.0, 0.1 M KCl, 2 mM MgCl_2 , 0.1 mM EDTA,

0.2 mM dithiothreitol (DTT), 0.1 mM phenylmethylsulphonyl fluoride (PMSF), 10% glycerol). The flow-through fraction was directly loaded onto a 5-ml Reactive green (RG) agarose column (Sigma) equilibrated with buffer H. The RG–1.0 M KCl fraction was dialysed against buffer A (20 mM HEPES pH 8.0, 0.1 M KCl, 0.2 mM EDTA, 0.2 mM DTT, 0.1 mM PMSF, 10% glycerol) and applied to a 5-ml Econo Q column (Bio-Rad) equilibrated with buffer A. Protein was eluted with a 90-ml linear gradient from 0.1 M to 1.0 M KCl (flow rate, 2 ml min⁻¹). The Econo Q fraction of the hSWI/SNF complex A (200 μ l) was fractionated on a Superose 6 HR 10/30 gel filtration column (Pharmacia) equilibrated with buffer A at 0.2 ml min⁻¹ and 0.6-ml fractions were collected. The Econo Q fraction of the hSWI/SNF complex B (440 μ l) was loaded onto a 15–40% linear glycerol gradient (5 ml) in buffer H. Centrifugation was at 35,000 r.p.m. for 12 h in an SW55 rotor (Beckman) and the gradient divided into 24 fractions. **d**, ATPase activity of the hSWI/SNF complex (300 ng, Superose 6 fraction) was measured in the presence or absence of 1 μ g of single-stranded pBluescript DNA or double-stranded pUC18 DNA as described²². The background hydrolysis (absence of hSWI/SNF complex) was 6 pmol and this value was subtracted. The results shown are average values of duplicate reactions.



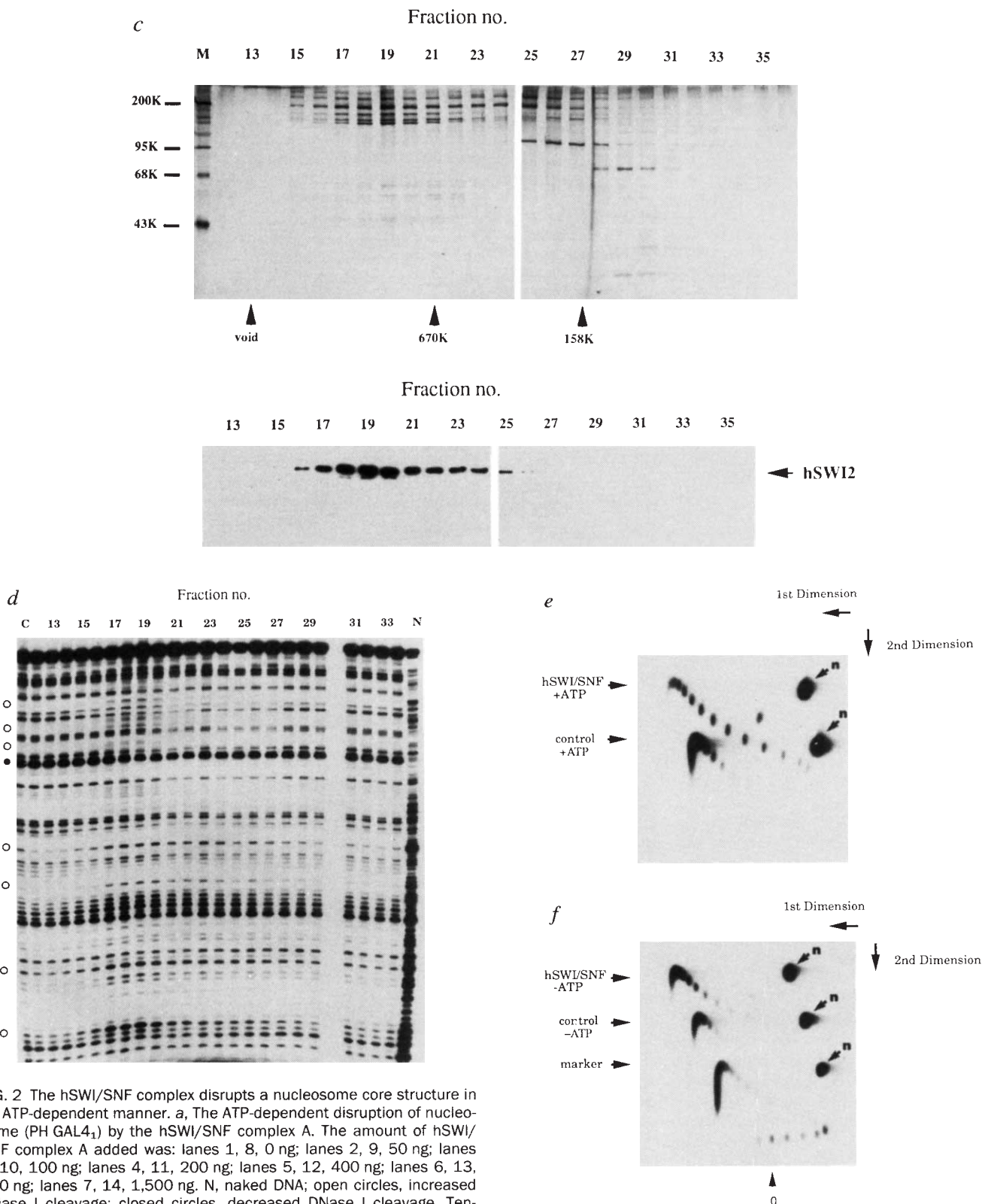


FIG. 2 The hSWI/SNF complex disrupts a nucleosome core structure in an ATP-dependent manner. **a**, The ATP-dependent disruption of nucleosome (PH GAL₄) by the hSWI/SNF complex A. The amount of hSWI/SNF complex A added was: lanes 1, 8, 0 ng; lanes 2, 9, 50 ng; lanes 3, 10, 100 ng; lanes 4, 11, 200 ng; lanes 5, 12, 400 ng; lanes 6, 13, 800 ng; lanes 7, 14, 1,500 ng. N, naked DNA; open circles, increased DNase I cleavage; closed circles, decreased DNase I cleavage. Ten-base-pair repeats calculated from the 5' end of the DNA fragment are indicated on the left. **b**, ATP hydrolysis requirement for nucleosome disruption. PH GAL₄ nucleosome was incubated with 1.3 µg of the hSWI/SNF complex A (left panel) and PH MLT (+3) nucleosome was incubated with 240 ng of the hSWI/SNF complex B (right panel). Arrows indicate modifications of DNase I digestion pattern similarly observed in both nucleosomes (PH GAL₄ and PH MLT (+3)). **c**, **d**, The hSWI/SNF complex and nucleosome-disrupting activity co-fractionate. hSWI/SNF complex A (Econo Q fraction) was fractionated by Superose 6 gel filtration chromatography, and each fraction was assayed for nucleosomal disruption activity. **c**, Upper part, Silver-stained SDS-PAGE gel; lower

part, immunoblotting assay with the α -BRG1 antibody, **d**, nucleosome disruption assay. Open circles, increased DNase I cleavage; closed circles, decreased DNase I cleavage. **e**, **f**, The hSWI/SNF complex decreases DNA superhelical density in reconstituted chromatin. DNA topoisomers were resolved by two-dimensional gel electrophoresis²³. The direction of electrophoresis in the first dimension (absence of chloroquine) and second dimension (presence of 4.1 µM chloroquine) are indicated. To avoid overlap, the samples were loaded in separate wells within a 1.5-h time interval. The nicked DNA (n) serves as a reference point for the migration of DNA.

METHODS. A 149-bp DNA containing two copies of a 20-bp artificial nucleosome positioning sequence¹² at the one end and a GAL4 binding site (PH GAL4₁) or adenovirus major late promoter TATA box (PH MLT (+3)) at the dyad position was reconstituted with purified HeLa histone octamers by a salt dilution method¹⁵. Glycerol gradient sedimentation revealed that >95% of the DNA was nucleosomal (data not shown). The integrity of the reconstituted nucleosome was confirmed by micrococcal nuclease digestion (data not shown; and see accompanying manuscript⁸). The reconstituted nucleosome (3.3 ng total DNA), isolated by a 5–30% glycerol gradient centrifugation (5 ml, 35,000 r.p.m., 15 h, Beckman SW55 rotor), was incubated in 25- μ l reaction buffer (12 mM HEPES pH 7.9, 60 mM KCl, 6 mM MgCl₂, 60 μ M EDTA, 2 mM DTT, 13% glycerol) with the hSWI/SNF complex for 30 min at 30 °C in the presence or the absence of 4 mM Mg-ATP, ATP- γ S or AMP-PNP. The nucleosome was digested with DNase I (0.32 U) for 2 min at room temperature, and the DNA products fractionated on an 8% polyacrylamide/8 M urea sequencing gel. e, f, pG₅HC₂AT (3.35 kb), internally labelled with ³²P, was reconstituted with HeLa histone octamers in a *Xenopus* oocyte heat-treated extract²⁴. Reconstituted chromatin was isolated on a 10–40% glycerol gradient (5 ml, 35,000 r.p.m., 4 h, SW55 Beckman rotor). The reconstituted chromatin (2.2 ng total DNA) in 25 μ l reaction buffer (12 mM HEPES pH 7.9, 60 mM KCl, 6 mM MgCl₂, 60 μ M EDTA, 4 mM Mg-ATP, 2 mM DTT, 13% glycerol, 3 units topoisomerase I (Promega)) was incubated at 30 °C for 40 min in the presence or absence (controls) of 900 ng of the hSWI/SNF complex. Two-dimensional gel electrophoresis was as previously described²³.

(hSWI2, p170, p150, p60, p50, p45 and p43) co-fractionated on a Superose 6 gel filtration column (hSWI/SNF complex A; Fig. 2c) and on a glycerol gradient (hSWI/SNF complex B; data not shown), consistent with the notion that they are components of a single complex with a native molecular mass greater than 700,000 (700 K). However, additional experiments will be required to define the subunit composition of the two hSWI/SNF complexes, to determine which polypeptides in these fractions are required for activity, and to ascertain whether there are functional differences between the two hSWI/SNF complexes. We note that in all the assays presented here, and in the accompanying manuscript⁸, the activities of the two hSWI/SNF complexes were similar.

Yeast SWI2 (ySWI2) and hSWI2 have a DNA helicase motif^{6,7,9,10} and indeed an *Escherichia coli*-derived ySWI2 has a DNA-dependent ATPase activity¹¹. Figure 1d shows that hSWI/SNF complex A contained an ATPase activity which was significantly stimulated by addition of DNA. Double-stranded DNA stimulated ATPase activity more effectively than single-stranded DNA, analogous to the results obtained with a recombinant ySWI2 protein¹¹.

We next examined the ability of the hSWI/SNF complex to alter nucleosome structure. A 149-base-pair (bp) DNA fragment containing two copies of a 20-bp artificial nucleosome positioning sequence¹² at one end and a single GAL4-binding site at its centre was reconstituted into a nucleosome using purified HeLa core histones. The DNase I digestion pattern of the reconstituted nucleosome showed the expected 10-bp periodicity (Fig. 2a, lane 1). In the absence of ATP, increasing amounts of the hSWI/SNF complex produced no qualitative change in the DNase I digestion pattern (Fig. 2a, lanes 2–7), but with ATP present it was dramatically changed; in particular, the characteristic 10-bp periodicity was progressively lost with increasing amounts of the hSWI/SNF complex (Fig. 2a, lanes 9–14). The complex had no effect on the DNase I digestion pattern of naked DNA (data not shown; and see accompanying manuscript⁸).

Figure 2b shows that the non-hydrolysable ATP analogues, AMP-PNP and ATP- γ S, did not support the nucleosome disrupting activity of the hSWI/SNF complex (left panel). This activity was not DNA-template-specific as the hSWI/SNF complex had a comparable effect on a nucleosome core containing the adenovirus major late promoter TATA box (Fig. 2b, right panel).

To obtain more evidence that the hSWI/SNF complex, and not a contaminant, caused the nucleosome disruption, partially purified hSWI/SNF complex A (Econo Q fraction) was fractionated on a Superose 6 column and the fractions were tested for their ability to disrupt nucleosome structure. Figure 2c shows that hSWI/SNF complex A, which included hSWI2, eluted from the Superose 6 column as a broad peak with a native molecular mass greater than 700K. Significantly, Fig. 2d shows that hSWI/SNF complex A and the nucleosome-disrupting activity co-fractionated, with a peak at fraction 19. Similarly, hSWI/SNF complex B co-fractionated with the nucleosome disrupting activity on a glycerol gradient (data not shown). Taken together, these results indicate that hSWI/SNF complexes A and B disrupt nucleosome structure in an ATP-dependent manner.

If nucleosome disruption is accompanied by either unwinding of nucleosomal DNA or removal of histones, the linking number per nucleosome will be altered. To test this possibility, ³²P-labelled closed circular DNA was reconstituted with purified histone octamers in a *Xenopus* oocyte heat-treated extract and the reconstituted chromatin was isolated by glycerol gradient centrifugation. On average, 15.5 nucleosomes were deposited on 3.35-kilobase (kb) DNA (one nucleosome per 216 bp of DNA). Figure 2e, f shows that following incubation with the hSWI/SNF complex, ATP and topoisomerase I, the average linking number of the isolated DNA was dramatically altered from 15.5 to 10.5, but in the absence of ATP, the complex had no effect

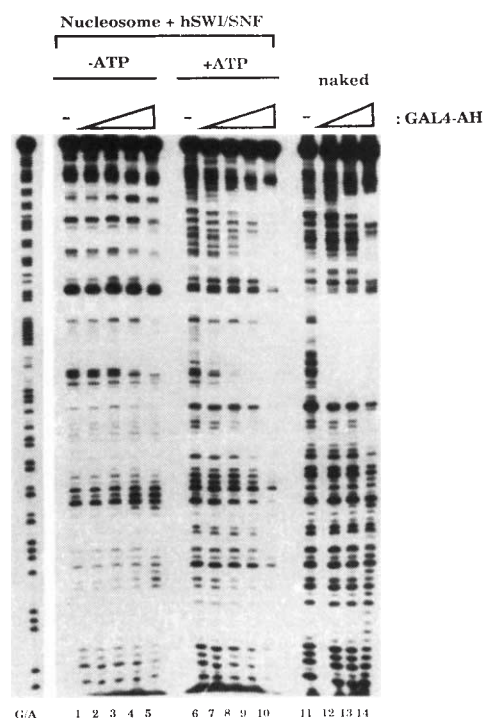


FIG. 3 The hSWI/SNF complex facilitates binding of GAL4-AH to nucleosomal DNA.

METHODS. The nucleosome core was reconstituted as described for Fig. 2. The reconstituted nucleosome (3.3 ng total DNA) was incubated in 25 μ l reaction buffer (12 mM HEPES pH 7.9, 60 mM KCl, 6 mM MgCl₂, 60 μ M EDTA, 2 mM DTT, 13% glycerol) with 1.3 μ g of the hSWI/SNF complex for 30 min at 30 °C in the presence or absence of 4 mM Mg-ATP. GAL4-AH was added followed by further incubation for 30 min at 30 °C. The nucleosome was digested with DNase I (0.4 U) for 2 min at room temperature, and analysed on an 8% polyacrylamide/8 M urea sequencing gel. The amount of GAL4-AH added was: lanes 1, 6, 11, 0 M; lanes 2, 7, 12, 2.4×10^{-8} M; lanes 3, 8, 13, 2.4×10^{-7} M; lanes 4, 9, 14, 2.4×10^{-6} M; lanes 5, 10, 2.4×10^{-5} M.

on the linking number of DNA isolated from reconstituted chromatin. These results further support the notion that the hSWI/SNF complex mediates the ATP-dependent disruption of a nucleosome core and that this disruption is accompanied by a dramatic change in DNA topology. In addition, our data indicate that the hSWI/SNF can alter the chromatin structure of a closed circular DNA template as well as a mono-nucleosome.

The observation that a hSWI/SNF complex disrupted a nucleosome raised the possibility that the hSWI/SNF complex could help transcription factors bind to their sites within a nucleosomal DNA template. To test this possibility, we first examined the effect of the hSWI/SNF complex on binding of the activator, GAL4-AH, to its DNA site located within a nucleosome core (Fig. 3). In the absence of the hSWI/SNF complex, GAL4-AH bound to its site within a nucleosome $\sim 10^4$ -fold less well than to naked DNA. In contrast, in the presence of both the hSWI/SNF complex and ATP, GAL4-AH binding was enhanced more than 100-fold. The hSWI/SNF complex had no effect without ATP.

We next investigated whether a transcriptional activation domain affected the ability of the hSWI/SNF complex to facilitate binding of a GAL4 derivative to a nucleosomal DNA template. Three GAL4 derivatives were analysed: GAL4(1-94), which lacks an activation domain; GAL4-AH, which has a relatively weak acidic activation domain¹³; and GAL4-VP16, which has a well-characterized, extremely potent acidic activation domain¹⁴. Previous studies have shown that on naked DNA or purified nucleosomal DNA templates, these activation domains do not affect GAL4 DNA binding (ref. 15; and data not shown). Table 1 shows that in three independent experiments GAL4-VP16 and GAL4-AH bound significantly better than GAL4(1-94) to nucleosomal DNA. Furthermore, binding of GAL4-VP16 was 5–10-fold higher than that of GAL4-AH.

In this and the accompanying manuscript⁸ we have shown that in the absence of promoter-specific and general transcription factors, a partially purified hSWI/SNF complex can disrupt the structure of a nucleosome core. On disruption of the nucleosome, binding of an activator and TATA-box binding protein (TBP) to nucleosomal DNA templates is augmented. The ySWI/SNF complex was originally identified as being required for the *in vivo* function of diverse activators^{4,16,17}. Our results strongly suggest that the SWI/SNF complex is required for activator function because it directly disrupts nucleosomal structure to increase access of transcription factors to promoters. The requirement for the SWI/SNF complex *in vivo* implies that disruption of chromatin structure is an essential aspect of transcriptional activation.

We have shown that the hSWI/SNF-complex-facilitated binding of three GAL4 derivatives parallels their ability to activate transcription. In particular, the VP16 activation domain was extremely active in this assay. Although the physical basis for this effect is unknown, the VP16 activation domain could interact with components of the hSWI/SNF complex or histones.

In vivo, GAL4-VP16 is a significantly more potent activator than GAL4-AH¹⁴. However, *in vitro* on naked DNA templates, GAL4-AH is a comparable, and in some instances an even better, activator than GAL4-VP16 (ref. 18). Thus, part of the high activity of the VP16 activation domain is lost on naked DNA templates. We speculate that the ability of the VP16 activation domain to enhance binding to nucleosomal DNA is part of the basis for its unusual potency. □

8. Imbalzano, A. N., Kwon, H., Green, M. R. & Kingston, R. E. *Nature* **370**, 481–485 (1994).
9. Laurent, B. C., Yang, X. & Carlson, M. *Molec. cell. Biol.* **12**, 1893–1902 (1992).
10. Henikoff, S. *Trends biochem. Sci.* **18**, 291–292 (1993).
11. Laurent, B. C., Treich, I. & Carlson, M. *Genes Dev.* **7**, 583–591 (1993).
12. Shrader, T. E. & Crothers, D. M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7418–7422 (1989).
13. Giniger, E. & Ptashne, M. *Nature* **330**, 670–672 (1987).
14. Sadowski, I., Ma, J., Triezenberg, S. & Ptashne, M. *Nature* **335**, 563–564 (1988).
15. Workman, J. L. & Kingston, R. E. *Science* **258**, 1780–1784 (1992).
16. Yoshinaga, S. K., Peterson, C. L., Herskowitz, I. & Yamamoto, K. R. *Science* **258**, 1598–1604 (1992).
17. Laurent, B. C. & Carlson, M. *Genes Dev.* **6**, 1707–1715 (1992).
18. Workman, J. L., Taylor, I. C. A. & Kingston, R. E. *Cell* **64**, 533–544 (1991).
19. Imbalzano, A. N., Zaret, K. S. & Kingston, R. E. *J. Biol. Chem.* **269**, 8280–8286 (1994).
20. Lin, Y. S., Carey, M. F., Ptashne, M. & Green, M. R. *Cell* **54**, 659–664 (1988).
21. Chasman, D. I., Leatherwood, J., Carey, M., Ptashne, M. & Kornberg, R. D. *Molec. cell. Biol.* **9**, 4746–4749 (1989).
22. Schwer, B. & Guthrie, C. *Nature* **349**, 494–499 (1991).
23. Peck, L. J. & Wang, J. C. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6206–6210 (1983).
24. Workman, J. L., Taylor, I. C. A., Kingston, R. E. & Roeder, R. G. *Meth. Cell Biol.* **35**, 419–447 (1991).

ACKNOWLEDGEMENTS. We thank L. L. Chiang for technical assistance; S. A. Brown for reagents; and C. Peterson for helpful discussion and communicating results before publication. This work was supported by grants from the NIH to M.R.G. and R.E.K.

Facilitated binding of TATA-binding protein to nucleosomal DNA

Anthony N. Imbalzano*, Hyockman Kwon†, Michael R. Green† & Robert E. Kingston*‡

* Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA, and Department of Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

† Howard Hughes Medical Institute, Program in Molecular Medicine, University of Massachusetts Medical Center, Worcester, Massachusetts 01605, USA

BINDING of the TATA-binding protein (TBP) to the TATA box is required for transcription from many eukaryotic promoters in gene expression. Regulation of this binding is therefore likely to be an important determinant of promoter activity. Incorporation of the TATA sequence into nucleosomes dramatically reduces transcription initiation^{1–3}, presumably because of stereochemical constraints on binding of general transcription factors. Biochemical and genetic studies imply that cellular factors such as yeast SWI/SNF are required for activator function and might alter chromatin structure^{4–11}. One step that could be regulated during the activation process is TBP binding in chromatin^{12,13}. We show here that binding of TBP to the TATA sequence is severely inhibited by incorporation of this sequence into a nucleosome. Inhibition can be overcome by ATP-dependent alterations in nucleosomal DNA structure mediated by hSWI/SNF, a putative human homologue of the yeast SWI/SNF complex. Additionally, the orientation of the TATA sequence relative to the surface of the histone core affects the access of TBP. We propose that the dynamic remodelling of chromatin structure to allow TBP binding is a key step in the regulation of eukaryotic gene expression.

We constructed a 149-base pair (bp) DNA fragment containing two rotational phasing sequences¹⁴ at the 5' end and the adenovirus major late promoter TATA box in the middle (PH MLT). The phasing sequence favours specific positioning of the histone core. Two variant templates positioned the TATA box either 3 or 6 bp closer to the phasing sequences (PH MLT(+3) and PH MLT(+6)), which should rotate the position of the TATA relative to the face of the histone core (see Fig. 1c and below). This should ensure that at least one of the templates contains the TATA sequences in an accessible position on the surface of the histones. *In vitro* assembled mononucleosomes¹⁵

‡ To whom correspondence should be addressed.

Received 4 February; accepted 7 July 1994.

1. Felsenfeld, G. *Nature* **355**, 219–224 (1992).
2. Taylor, I. C. A., Workman, J. L., Schuetz, T. J. & Kingston, R. E. *Genes Dev.* **5**, 1285–1298 (1991).
3. Pina, B., Bruggemeier, U. & Beato, M. *Cell* **60**, 719–731 (1990).
4. Peterson, C. L. & Herskowitz, I. *Cell* **68**, 573–583 (1992).
5. Hirschhorn, J. N., Brown, S. A., Clark, C. D. & Winston, F. *Genes Dev.* **6**, 2288–2298 (1992).
6. Muchardt, C. & Yaniv, M. *EMBO J.* **12**, 4279–4290 (1993).
7. Khavari, P. A., Peterson, C. L., Tamkun, J. W., Mendel, D. B. & Crabtree, G. R. *Nature* **366**, 170–174 (1993).

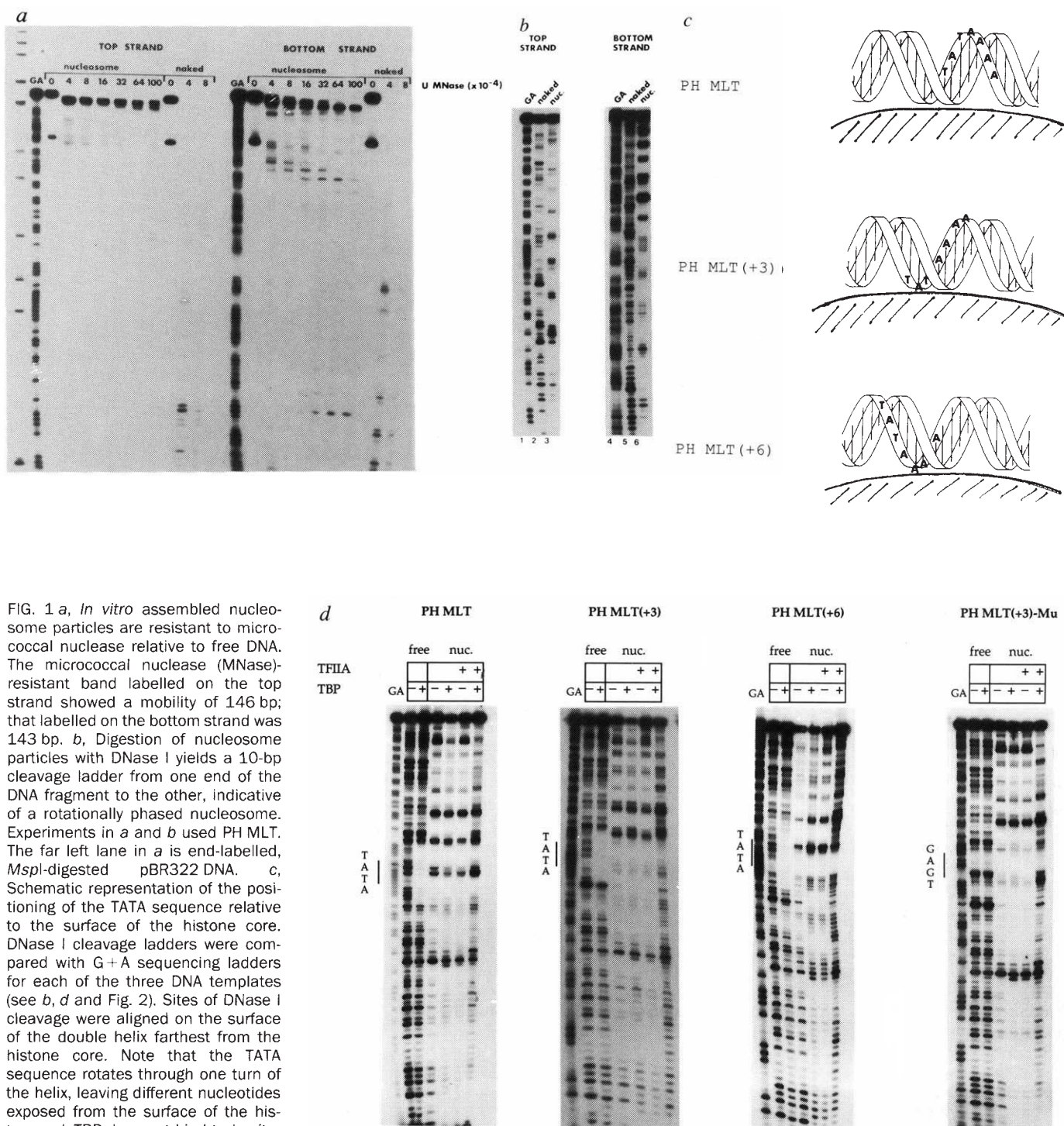


FIG. 1 *a*, *In vitro* assembled nucleosome particles are resistant to micrococcal nuclease relative to free DNA. The micrococcal nuclease (MNase)-resistant band labelled on the top strand showed a mobility of 146 bp; that labelled on the bottom strand was 143 bp. *b*, Digestion of nucleosome particles with DNase I yields a 10-bp cleavage ladder from one end of the DNA fragment to the other, indicative of a rotationally phased nucleosome. Experiments in *a* and *b* used PH MLT. The far left lane in *a* is end-labelled, *Msp*I-digested pBR322 DNA. *c*, Schematic representation of the positioning of the TATA sequence relative to the surface of the histone core. DNase I cleavage ladders were compared with G+A sequencing ladders for each of the three DNA templates (see *b*, *d* and Fig. 2). Sites of DNase I cleavage were aligned on the surface of the double helix farthest from the histone core. Note that the TATA sequence rotates through one turn of the helix, leaving different nucleotides exposed from the surface of the histones. *d*, TBP does not bind to *in vitro* assembled mononucleosomes containing a TATA box at the dyad axis of symmetry, alone or in the presence of TFIIA.

METHODS. The PH MLT restriction fragment²⁹ contains two copies of the 20-bp 'GT' artificial nucleosome phasing sequence¹⁴ at the 5' end and the adenovirus major late promoter TATA box in the centre of the fragment. Incorporation into a nucleosome places the TATA at the dyad axis of symmetry. To create the (+3) and (+6) variants, double-stranded oligonucleotides containing the TATA sequence 3 or 6 bp closer to the phasing sequences were subcloned into *Bgl*II/*Nco*I-digested PH MLT²⁹. The mutant PH MLT (+3) template is identical to the (+3) variant except the sequence GAGTCCC replaced the TATA sequence. Nucleosomes were assembled *in vitro* by salt dilution¹⁵ using 450 ng of end-labelled fragment, 5 μ g of *Hae*III-digested pUC18, and 8.6 μ g of core histone octamers, purified from HeLa cells as described³⁰, and were then purified on glycerol gradients. For analysis, 0.3 ng of labelled nucleosome

or free DNA (1.2×10^{-10} M) was combined in a 25- μ l reaction containing 12 mM HEPES pH 7.9, 60 mM KCl, 7 mM MgCl₂, 15% glycerol, 0.6 mM dithiothreitol, 0.06 mM EDTA and 500 ng BSA. Reactions with free DNA also contained 100 ng poly-dGdC (Pharmacia). Where indicated, binding reactions contained 1.5×10^{-6} M TBP and 2×10^{-6} M TFIIA, except for naked DNA samples, which contained 2×10^{-8} M TBP. Recombinant yeast TBP and recombinant yeast TFIIA preparations were characterized and described as previously²¹. Reactions were incubated at 30 °C for 25 min, then treated with MNase (a) or with 0.02 U (free DNA, *b*, *d*), or 0.2 U (nucleosomal DNA, *b*, *d*) of RQ1 DNase I (Promega) for 2 min at room temperature. Samples were prepared for electrophoresis as described²¹.

were resistant to micrococcal nuclease digestion (Fig. 1a) and, on DNase I digestion, yielded a 10-bp 'ladder' which spanned the length of the fragment, typical of rotationally phased nucleosomes (Fig. 1b). The 10-bp ladder allowed the determination of the orientation of the TATA sequence relative to the surface of the histone core (Fig. 1c).

Nucleosomes were then incubated with recombinant *Saccharomyces cerevisiae* TBP and/or transcription factor IIA (TFIIA). TBP is the DNA-binding component of TFIID, which is the first factor to bind during formation of the preinitiation complex (PIC) on eukaryotic RNA polymerase II promoters^{16,17}. Prebinding of either TFIID or TBP protects the template from repression by subsequent assembly of nucleosomes^{2,12,13}. Many transcriptional activators show increased levels of stimulation on nucleosomal templates^{4,18,20}, even when TBP is substituted for TFIID^{12,13}, suggesting that some activators can enhance the ability of the PIC to form in chromatin and that mechanisms that facilitate TBP binding to nucleosomal DNA do not absolutely require the TBP-associated factors that also comprise TFIID activity. General factor TFIIA was tested in combination with TBP, as TFIIA can modulate binding of TBP^{16,17,21}.

DNase I footprinting showed no indication of specific binding to any of the three nucleosomal templates, even at TBP concentrations as high as 4×10^{-5} M, which is four orders of magnitude greater than required for binding to naked DNA^{22,23} (Fig. 1d,

and data not shown). When TBP and TFIIA were both present, again there was no specific protection over the TATA box, although an increased number of DNase I cleavages occurred on all three templates, resulting in a partial loss of the 10-bp ladder. A template containing a mutated TATA box showed a similar DNase I cleavage pattern (Fig. 1d, far right), so the increased nuclease sensitivity caused by TBP and TFIIA was not due to binding to a TATA sequence. Nucleosome formation therefore dramatically inhibited TBP binding, suggesting that repression of basal transcription by nucleosomes is due to an inability of the PIC to form.

We next examined possible mechanisms which might be used to alter nucleosome structure and facilitate interaction by TBP. When nucleosomes were incubated with a fraction containing the hSWI/SNF complex²⁴ and then treated with DNase I, an ATP-dependent disruption of the nucleosomal 10-bp ladder was observed such that the DNase I cleavage pattern resembled but was not identical to that of naked DNA (Fig. 2a; compare lanes 8–10 with 11 and 12). Electrophoresis of nucleosomes treated with hSWI/SNF plus ATP on a native polyacrylamide gel showed that their mobility was identical to that of untreated nucleosomes, suggesting that the histones were still associated with the DNA (data not shown). The ability of hSWI/SNF to alter accessibility of the nucleosomal DNA to DNase suggested that it might also increase the accessibility of TBP. When TBP or TBP plus TFIIA were added after the nucleosome had been

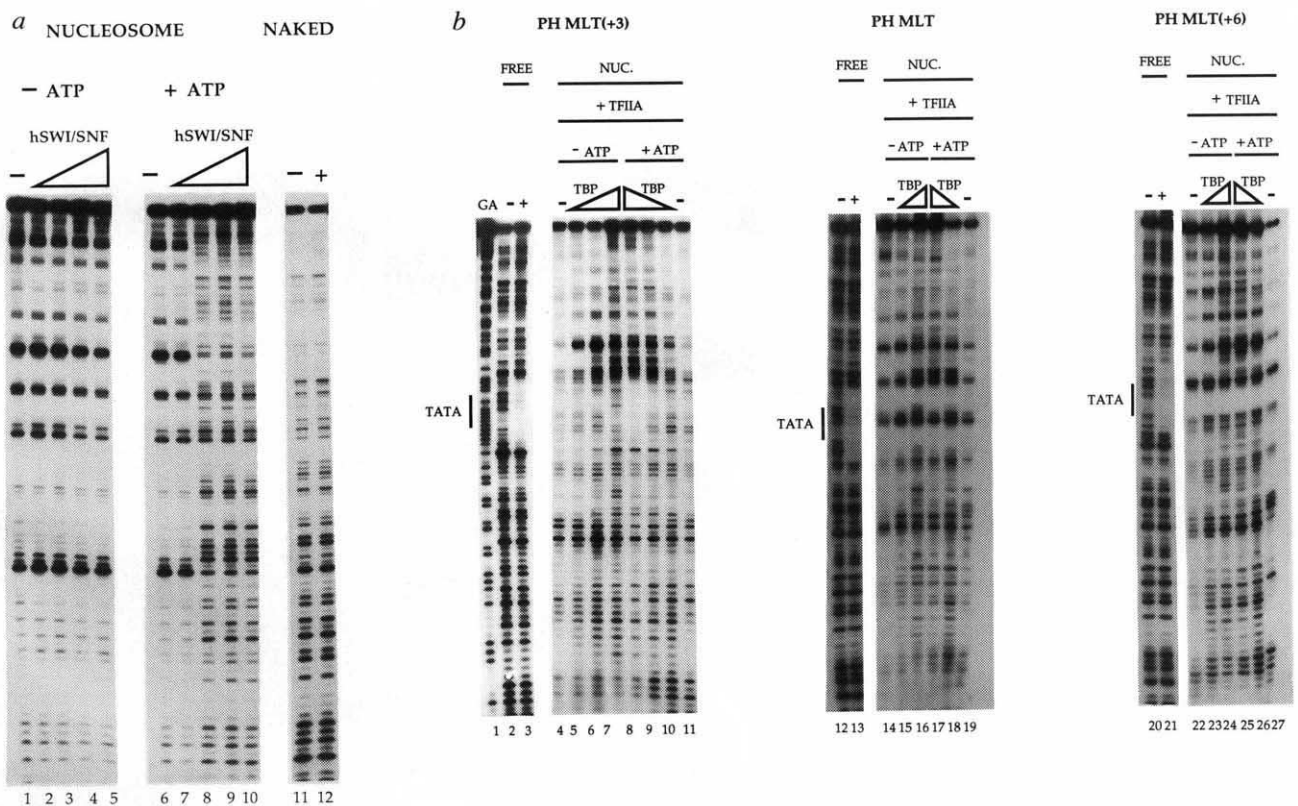


FIG. 2 a, hSWI/SNF alters nucleosomal DNA structure in an ATP-dependent manner. b, Prior exposure of nucleosomes to hSWI/SNF plus ATP allows binding of TBP in the presence of TFIIA to PH MLT(+3) nucleosomes but not to PH MLT or PH MLT(+6) nucleosomes.

METHODS. a, Binding reactions contained PH MLT(+3) nucleosomal DNA and were done as described for Fig. 1 in the presence of increasing amounts of hSWI/SNF (lanes 2 and 7, 12 ng; lanes 3 and 8, 120 ng; lanes 4 and 9, 240 ng; lanes 5, 10 and 12, 480 ng); Econo Q fraction derived from the phosphocellulose 0.7 M cut²⁴ in the presence or absence of 4 mM Mg-ATP. Samples were then treated with DNase I as described for Fig. 1. Identical results were obtained when we used the

Econo Q fraction containing hSWI/SNF derived from the phosphocellulose 0.5 M cut²⁴. b, Reactions were done as described in a. After incubation with hSWI/SNF (240 ng), TBP or TBP + TFIIA were added and the reactions were incubated at 30 °C for another 25 min before treatment with DNase I and subsequent electrophoresis. For clarity, only lanes with TFIIA are shown. TBP concentrations were 1.5×10^{-8} M (lanes 5, 10), 1.5×10^{-7} M (lanes 6, 9, 15, 18, 23, 26) and 1.5×10^{-6} M (lanes 7, 8, 16, 17, 24, 25). TFIIA was present at 2×10^{-6} M. As in a, identical results were obtained using the Econo Q fraction containing hSWI/SNF derived from the phosphocellulose 0.5 M cut²⁴.

TABLE 1 TBP concentration required for half-maximal protection of TATA sequence

		– ATP	+ ATP
Naked DNA	TBP	2×10^{-9} M*	2×10^{-9} M
	TBP + TFIIA	2×10^{-9} M*	1×10^{-9} M
	TBP + hSWI/SNF	2×10^{-9}	2×10^{-8}
	TBP + TFIIA + hSWI/SNF	2×10^{-9}	2×10^{-9}
Nucleosomal DNA	TBP	$> 4 \times 10^{-5}$ M	Not done
	TBP + TFIIA	$> 4 \times 10^{-5}$	$> 2 \times 10^{-6}$
	TBP + hSWI/SNF	$> 2 \times 10^{-6}$	$> 2 \times 10^{-6}$
	TBP + TFIIA + hSWI/SNF	$> 2 \times 10^{-6}$	1×10^{-7}

TBP concentrations required for half-maximal protection of naked and nucleosomal PH MLT(+3) DNA were determined in the presence and absence of TFIIA, hSWI/SNF and ATP. Reactions were as described for Fig. 1. Where indicated, reactions contained 240 ng hSWI/SNF and 2×10^{-6} M TFIIA. TBP concentrations required for half-maximal footprinting were determined as previously described²¹.

* Data from ref. 21.

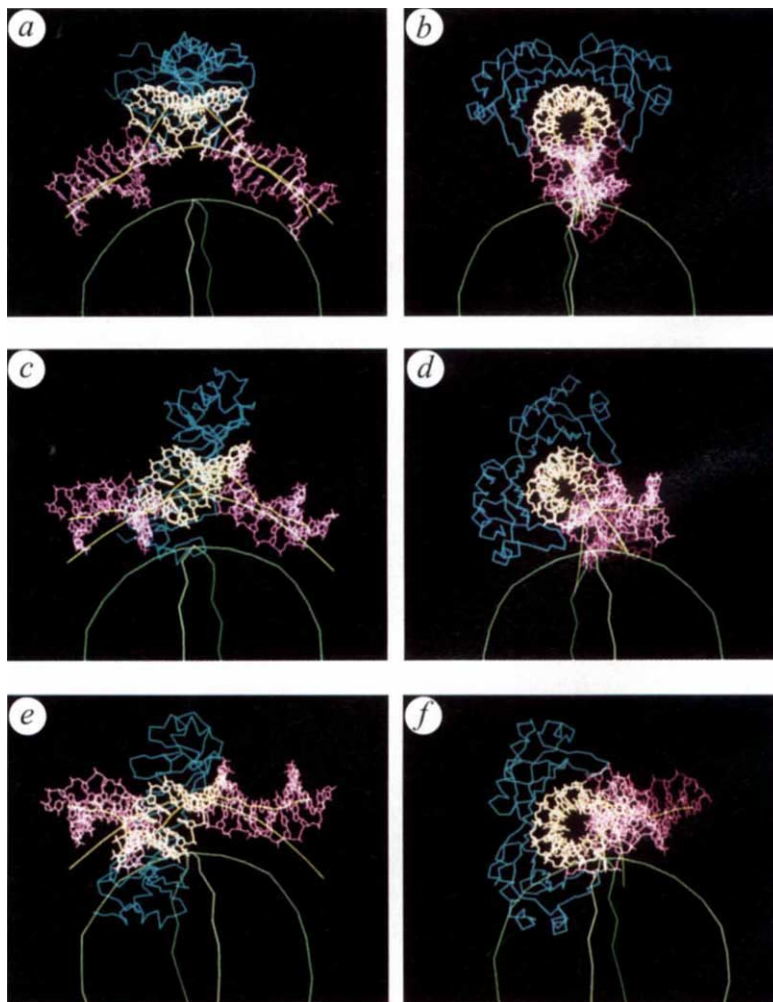
incubated with hSWI/SNF, ATP-dependent protection of the TATA box occurred on PH MLT(+3) only when both TBP and TFIIA were present (Fig. 2b; compare lanes 7 and 8; Table 1; and data not shown). Thus, hSWI/SNF sufficiently alters nucleosomal DNA to facilitate specific binding to the TATA box of PH MLT(+3). No protection of the TATA box was observed on templates PH MLT or PH MLT(+6) (Fig. 2b, lanes 16, 17, 24, 25), showing that only one orientation of the TATA box relative to the histone core was rendered accessible to TBP

plus TFIIA. This indicates that only specific surface orientations of the TATA box are bound stably by TBP and TFIIA under these conditions, and therefore implies that an effect of nucleosomal structure on TBP binding persists even in the presence of hSWI/SNF.

hSWI/SNF did not alter the DNase I digestion pattern of naked DNA (data not shown). However, hSWI/SNF, in the presence of ATP, inhibited binding of TBP to naked DNA by 10-fold (Table 1). The presence of TFIIA counteracted the inhibitory effect of the hSWI/SNF fraction on TBP binding to naked DNA. This may explain the requirement for TFIIA in hSWI/SNF-facilitated binding to nucleosomal DNA.

The observation that TBP binding depends on the rotational position of the TATA sequence has no precedent among tested transcription factor:nucleosome interactions. We have found that TBP binding to nucleosomal templates can also be facilitated in the same rotationally dependent manner by including hyperacetylated histones in the nucleosome assembly (A.N.I. and R.E.K., unpublished observations). Therefore, the dependence on rotational positioning for TBP binding is not specific to hSWI/SNF activity. These findings are perplexing, given that the TATA sequence of PH MLT(+3) has the most contact between the minor groove and the histone core of the three orientations (see Fig. 1c) and that TBP has been shown to bind to DNA in the minor groove²⁵. In an attempt to understand these results, we have modelled the TBP–nucleosome interaction using the recently solved TBP:DNA co-crystal structure^{26,27} (Fig. 3; modelling was done with Y. Kim and P. Sigler). A skeletal representation of yeast TBP bound to a consensus TATA

FIG. 3 Skeletal models of recombinant yeast TBP (blue; α trace) bound to a consensus TATA box (white) that is extended in both directions by one turn of B-form DNA (pink) juxtaposed onto a ball (green) of appropriate size (~ 65 Å) for a histone core particle. Orientation of the TBP:TATA complex was based on assigned histone–DNA contacts made by a double helix on the globular histone core particle (depicted in Fig. 1c and derived from DNase cleavage ladders; see Figs 1 and 2). The yellow line represents the axis of DNA as it exists in the models. The green line represents the axis of B-form DNA that would exist if the DNA were juxtaposed against the globular core. a, c, e, Models of TBP bound to PH MLT, PH MLT(+3), or PH MLT(+6) nucleosomes, respectively. b, d, f, Models depicted in a, c and e, respectively, rotated by 90° around the vertical axis. Although the representation of the nucleosome in these models is spherical, the core of the histone octamer more closely resembles a cylinder. Precise modelling on a cylinder will require determination of the structure and path of the DNA around the histone core; however, a more precise representation of TBP:nucleosome interactions is unlikely to change the conclusions based on the models presented here.



sequence was aligned to a ball of the appropriate size for a histone core. Note that the 10-bp periodicity of DNase cleavage outside the footprint did not change on nucleosomes that were bound by TBP (Fig. 2*b*), as expected because the gross structural change caused by TBP binding to DNA is topologically neutral^{26,27}. This allowed us to align the DNA helix rotationally on the nucleosome using the assigned histone-DNA contacts derived from DNase cleavage ladders (Fig. 1*c*). These models are limited by the facts that no high-resolution structures are available for the nucleosome and that existing structures depicting the histone core lack the amino-terminal portions of the histones, which presumably extend from the globular core domain.

Construction of a model of TBP bound to the TATA box as positioned in PH MLT, which does not bind to TBP, indicates no steric hindrance. However, the bend in DNA caused by TBP has a much smaller radius of curvature (25 Å; refs 25, 26) than is expected for the bend of nucleosomal DNA according to current models (average radius of ~42 Å (ref. 28); see Fig. 3*a* and *b*). This difference in the radius of bending requires that the TATA sequence lifts away from the histone core when TBP is bound to PH MLT, resulting in the loss of histone-DNA contacts. In contrast, the TBP molecule on template PH MLT(+3) (where TBP can bind) is located such that the bend in DNA induced by TBP is tangential to the surface of the nucleosome, and the empty space depicted in the PH MLT model is not evident (Fig. 3*c* and *d*). Although the direction of the DNA extending from the TBP-TATA complex would need to be altered on this template so as to wrap around the histones, DNA is normally bent by association with the histone core. Finally, positioning TBP

on the TATA sequence oriented in PH MLT(+6), which does not bind TBP, indicates significant steric hindrance (Fig. 3*e* and *f*). These models therefore appear consistent with the experimental results, and indicate that two separate protein-DNA interactions that both bend DNA can be accommodated in a single complex if one bend is tangential to the surface of the second structure (see Fig. 3*c* and *d*).

These data could suggest that TBP and the general factors can only gain access to promoters that have a specifically positioned TATA sequence in intact cells. We feel that this is possible, but unlikely. Instead, we suggest that the data presented here reflect only the contributions of a few of the functional components that would be required *in vivo*. Here we observed binding to a specifically positioned TATA box only after alteration of nucleosome DNA structure by hSWI/SNF activity at TBP concentrations in the 10⁻⁷ M range, which is 100-fold more than is required to bind naked DNA^{22,23}. The presence of other cellular factors such as TBP-associated factors and gene-specific activator proteins might further increase binding of TBP.

The facilitated binding of TBP to nucleosomal DNA by hSWI/SNF (Fig. 2*b*) suggests a possible mechanism for overcoming nucleosome-mediated repression of general transcription factor activity. As the histones are apparently not removed by hSWI/SNF activity and as hSWI/SNF alters the linking number of nucleosomal DNA²⁴, hSWI/SNF activity might lower the activation energy for recognition of the TATA sequence by making the relatively stable and static contacts between DNA and the core histones less stable, and hence more accessible. This proposed ability to alter DNA-histone contacts may play a key role in the activation of promoters in chromatin. □

Received 2 February; accepted 7 July 1994.

1. Knezetic, J. A. & Luse, D. S. *Cell* **45**, 95-104 (1986).
2. Workman, J. L. & Roeder, R. G. *Cell* **51**, 613-622 (1987).
3. Matsui, T. *Molec. cell. Biol.* **7**, 1401-1408 (1987).
4. Lorch, Y., LaPointe, J. W. & Kornberg, R. D. *Genes Dev.* **6**, 2282-2287 (1992).
5. Crosten, G. E., Laybourn, P. J., Paranjape, S. M. & Kadonaga, J. T. *Genes Dev.* **6**, 2270-2281 (1992).
6. Peterson, C. L. & Herskowitz, I. *Cell* **68**, 573-583 (1992).
7. Laurent, B. C. & Carlson, M. *Genes Dev.* **6**, 1707-1715 (1992).
8. Yoshinaga, S. K., Peterson, C. L., Herskowitz, I. & Yamamoto, K. R. *Science* **258**, 1598-1604 (1992).
9. Hirschhorn, J. N., Brown, S. A., Clark, C. D. & Winston, F. *Genes Dev.* **6**, 2288-2298 (1992).
10. Mataliana, E., Franco, L. & Perez-Ortin, J. E. *Molec. gen. Genet.* **231**, 395-400 (1992).
11. Winston, F. & Carlson, M. *Trends Genet.* **8**, 387-391 (1992).
12. Meisterernst, M., Horikoshi, M. & Roeder, R. G. *Proc. natn. Acad. Sci. U.S.A.* **87**, 9153-9157 (1990).
13. Workman, J. L., Taylor, I. C. A. & Kingston, R. E. *Cell* **64**, 533-544 (1991).
14. Shrader, T. E. & Crothers, D. M. *Proc. natn. Acad. Sci. U.S.A.* **86**, 7418-7422 (1989).
15. Perlmann, T. & Wrangé, O. *EMBO J.* **7**, 3073-3079 (1988).
16. Buratowski, S., Hahn, S., Guarente, L. & Sharp, P. A. *Cell* **56**, 549-561 (1989).
17. Maldonado, E., Ha, I., Cortes, P., Weis, L. & Reinberg, D. *Molec. cell. Biol.* **10**, 6335-6347 (1990).
18. Workman, J. L., Abmayer, S. M., Cromlish, W. A. & Roeder, R. G. *Cell* **55**, 211-219 (1988).

19. Workman, J. L., Roeder, R. G. & Kingston, R. E. *EMBO J.* **9**, 1299-1308 (1990).
20. Kamakaka, R. T., Bulger, M. & Kadonaga, J. T. *Genes Dev.* **7**, 1779-1795 (1993).
21. Imbalzano, A. N., Zaret, K. S. & Kingston, R. E. *J. biol. Chem.* **269**, 8280-8286 (1994).
22. Hahn, S., Buratowski, S., Sharp, P. A. & Guarente, L. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5718-5722 (1989).
23. Hoopes, B. C., LeBlanc, J. F. & Hawley, D. K. *J. biol. Chem.* **267**, 11539-11547 (1992).
24. Kwon, H., Imbalzano, A. N., Khavari, P. A., Kingston, R. E. & Green, M. R. *Nature* **370**, 477-481 (1994).
25. Starr, D. B. & Hawley, D. K. *Cell* **67**, 1231-1240 (1991).
26. Kim, Y., Geiger, J. F., Hahn, S. & Sigler, P. B. *Nature* **365**, 512-520 (1993).
27. Kim, J. L., Nikolov, D. B. & Burley, S. K. *Nature* **365**, 520-527 (1993).
28. Arents, G. & Moudrianakis, E. N. *Proc. natn. Acad. Sci. U.S.A.* **90**, 10489-10493 (1993).
29. Wechsler, D. S., Papoulas, O., Dang, C. V. & Kingston, R. E. *Molec. cell. Biol.* **14**, 4097-4107 (1994).
30. Workman, J. L., Taylor, I. C. A., Kingston, R. E. & Roeder, R. G. *Meth. Cell Biol.* **35**, 419-447 (1991).

ACKNOWLEDGEMENTS. We thank P. Khavari and G. Crabtree for reagents needed to purify hSWI/SNF; Y. Kim and P. Sigler for help with the TBP modelling; A. Wolffe and J. Workman for discussion; and F. Winston, K. Struhl, M. Oettinger, K. Zaret, C. Bunker and J. Yodh for reviewing the manuscript. This work was supported by grants from the NIH to R.E.K. A.N.I. is supported by an NIH postdoctoral fellowship.

Let's get physical

A complete physical map of the region encompassing the locus for hereditary breast and ovarian cancer should assist the long-awaited cloning of this important gene.

FOR the better part of four years now, researchers the world over have been scouring a steadily diminishing stretch of the long arm of chromosome 17 for an elusive gene, *BRCA1*, responsible for hereditary breast and ovarian cancer. Women who inherit a faulty copy of *BRCA1* have about an 85 per cent risk of developing breast cancer during their lifetime, and together with one or more other loci, *BRCA1* accounts for about 5 per cent of all cases of breast cancer. Given that there are hundreds of thousands of new cases diagnosed around the world each year, the implications for screening families at high risk, and learning more about the molecular basis of both inherited and sporadic forms of the disease, place *BRCA1* firmly at the top of the human geneticist's wanted list.

Yet *BRCA1* is stubbornly refusing to reveal itself. For the past three years, researchers have been studying hundreds of families with inherited breast (and ovarian) cancer to refine the genetic linkage first detected at the end of 1990 (ref. 1). Thanks in large measure to an international consortium of researchers², a handful of important meiotic recombination events have been characterized, which in turn have allowed investigators to position *BRCA1* on one side or the other of a particular DNA marker, each time pushing in the boundaries of the so-called critical region. In parallel with these efforts, several groups have been establishing detailed physical maps using yeast artificial chromosomes (YAC), P1 clones and so on, to allow them to begin the tedious task of isolating the dozens of transcripts in the crucial region that should include *BRCA1* itself.

A paper in this month's *Nature Genetics*³ affords a glimpse at the strategies being employed jointly by two of the larger groups involved in this effort. A little more than two years ago, Ray White in Salt Lake City and Bruce Ponder in Cambridge agreed to pool their resources in their quest for the *BRCA1* gene³. The physical map they have

constructed consists of more than 130 different clones and embraces more than 3.5 megabases of genomic DNA. As is customary with ventures of this sort, the map is a little tentative in a couple of places (because of gaps or possible YAC rearrangements) but fortunately these are not in the area that really matters. By mapping the two closest (published) DNA markers flanking *BRCA1* that have been shown to be recombinant in key breast cancer families (*D17S78* and *D17S776*)^{4,5}, they estimate that the interval containing the gene spans no more than 1–2 centimorgans — about one megabase or so.

With classical genetics having just about exhausted its usefulness in localizing *BRCA1*, the next task is to trawl across the region in search of transcripts. Albertsen *et al.*³ pulled out dozens of complementary DNA clones by screening a fetal cDNA library with purified inserts from a pair of critical YACs, although a number of other techniques, such as exon trapping, are equally valid. Among the new genes described in the current report are relatives of the Rab⁵ family, the Ki antigen and a yeast transcriptional factor, GCN5. A rapid screen for mutations in samples from breast cancer patients has excluded some of these sequences, but the search continues apace.

Several other groups have similar physical and transcriptional maps at their disposal, and as more and more candidates are investigated⁶ it should not be long before the real culprit is nabbed. Whether the gene will be of immediate use in revealing the cellular process affected in (hereditary) breast cancer remains to be seen, for it is a distinct possibility that *BRCA1* will look quite unlike anything characterized in the public databases. Such has been the disheartening experience for a number of significant genetic discoveries recently, including the defect responsible for Huntington's disease⁷, polycystic kidney disease⁸ and spinocerebellar ataxia type 1 (see Banfi *et al.*⁹ in this month's issue of *Nature Genetics*). On the other hand, perusal of cDNA sequence databases has paid off handsomely with the identification of genes for hereditary non-polyposis colon cancer. And one should not forget that conventional biochemistry can still teach us a trick or two, as was nicely demonstrated just a few weeks ago with the revelation that the gene underlying Miller-Diecker lissencephaly (isolated

last year by positional cloning) is that for a subunit of the platelet-activating factor acetylhydrolase in brain¹⁰.

Physical mapping also plays an illuminating role in another paper in this month's issue¹¹, here in the search for a curious gene on the short arm of the X chromosome (Xp) which has a profound effect on the route of sexual development. In males, the presence of the *SRY* sex-determining gene (on the Y chromosome) is normally sufficient to initiate the development of the male gonads¹². But in rare individuals with an intact Y chromosome and partial duplications of the short arm of the X, female (or ambiguous) genitalia may develop.

Giovanna Camerino, from the University of Pavia in Italy, and colleagues¹¹ have studied eight patients with Xp duplications, four of whom showed sex reversal. They found that the duplications in the four sex reversed patients all had different breakpoints, making it most unlikely that their phenotype resulted from the disruption of a single gene. What they had in common, however, was the duplication of a small region estimated to be no larger than 160 kilobases. Presumably, the extra copy of a gene(s) in this interval, which Camerino calls *DSS* (for dosage-sensitive sex reversal), upsets normal testis formation, although interestingly the extent of sex reversal (gonadal dysgenesis) varied among the four patients. By contrast, deletion of this same region in patients with an otherwise normal male (46, XY) genotype does not affect the development of male genitalia, prompting Camerino and colleagues to speculate that *DSS* may represent a link between the ovarian and testicular developmental pathways.

Kevin Davies

Kevin Davies is the editor of Nature Genetics.

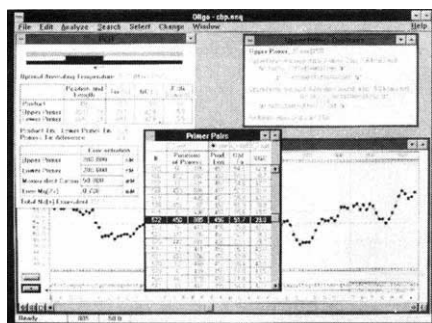
Also in this month's *Nature Genetics*: The genes involved in the t(X;18) translocation in synovial sarcoma; interaction of the *Rb* and *p53* genes; a keratin 2e mutational hotspot in ichthyosis bullosa of Siemens; and modelling the expansion of the Huntington's disease gene.

1. Hall, J.M. *et al.* *Science* **250**, 1684–1689 (1990).
2. Easton, D.F. *et al.* *Am. J. hum. Genet.* **52**, 678–701 (1993).
3. Albertsen, H. *et al.* *Nature Genet.* **7**, 472–479 (1994).
4. Simard, J. *et al.* *Hum. molec. Genet.* **2**, 1193–1199 (1993).
5. Goldgar, D.E. *et al.* *J. natn. Cancer Inst.* **86**, 200–209 (1994).
6. Brown, M.A. & Solomon, E. *Curr. Opin. Genet. Dev.* **4**, 439–445 (1994).
7. The Huntington's Disease Collaborative Research Group *Cell* **72**, 371–383 (1993).
8. The European Polycystic Kidney Disease Consortium *Cell* **77**, 881–894 (1994).
9. Banfi, S. *et al.* *Nature Genet.* **7**, 513–520 (1994).
10. Hattori, M. *et al.* *Nature* **370**, 216–218 (1994).
11. Bardoni, B. *et al.* *Nature Genet.* **7**, 497–501 (1994).
12. Nordqvist, K. & Lovell-Badge, R. *Nature Genet.* **7**, 7–9 (1994).

Peptide and oligo primer

New in the marketplace — the latest upgrade to a primer analysis software program, IL-1 β converting enzyme substrates, peptide and oligonucleotide synthesizers and assorted purification systems.

VERSION 5.0 of OLIGO, **primer analysis software** for Windows from National Biosciences, is now shipping (*Reader Service No. 100*). This latest Windows version includes several new functions to assist researchers in the selection of optimal oligonucleotides for PCR, sequencing, hybridizations and other applications. New features



NBI's primer analysis program upgrade.

include an oligonucleotide database that permits porting of oligonucleotide sequence and data to and from the program. The database saves oligo data pertinent to synthesis, and generates a convenient synthesis order form. The new primer search function finds all cross-compatible primers in a sequence file and orders them by either upper primer position number, or by PCR product size. Within this feature is a 5' GC clamp and automatic multiplex primer selections. NBI has also included an automatic relaxation of the search parameters if the user-defined search does not find primers at the original settings. OLIGO 5.0 also features automatic degenerate primer selection from an amino acid sequence, and permits back translation from any one of a comprehensive list of codon preference tables, or a user-defined table. Hybridization time calculations, an automatic search for ligase chain reaction primers, batch searching for unique oligos in multiple files, and restriction enzyme searches are also featured.

ω -agatoxin IVa is a peptide isolated from the venom of the American funnel web spider, *Agelenopsis aperta*. It has been shown to be a selective and potent **blocker of the mammalian P-type voltage-dependent calcium channel**. The Swiss Bachem Group is now making this product available to the research community upon request (*Reader Service No. 101*). IL-1 β converting enzyme (ICE) is an important mediator of inflammation. Bachem offers a peptide substrate that contains the protease cleavage site situated between two fluorophores located at the

termini of the molecule. Upon cleavage of this substrate, DABCYL-Tyr-Val-Ala-Asp-Ala-Pro-Val-EDANS (**IL-1 β converting enzyme substrate II**), Bachem says that an increase in fluorescence is observed at the EDANS emission wavelength of 490 nm, permitting a continuous assay of ICE that is useful in the screening of inhibitory compounds. **IL-1 β converting enzyme substrate I**, a 14-amino acid peptide that corresponds to the native cleavage site sequence around Asp¹¹⁶-Ala¹¹⁷ and inhibits the cleavage of the precursor to the mature protein, is also available from Bachem.

Novex has introduced precast 6 per cent TBE-urea **denaturing polyacrylamide gels**, which are intended for RNase protection assays, *in vitro* transcription studies and the analysis of large DNA oligos (*Reader Service No. 102*). The gels contain 7 M urea and provide good resolution of RNA fragments from 50 to 600 bases, says Novex. The gels can be stained with ethidium bromide or silver stain, or can be used for autoradiography, ultraviolet shadowing or northern blots. All products are said to be RNase- and DNase-free. Premixed sample and running buffers are also available. In addition to the 6 per cent gels, Novex offers precast 15 and 10 per cent TBE-urea gels for the analysis or purification of synthetic DNA or RNA oligos from 20 to 130 bases in length. These, and the 6 per cent gels are available in 1-mm thickness with 10 or 15 wells, and 1.5-mm thickness with a two-dimensional/preparative well format.

■ Synthesis

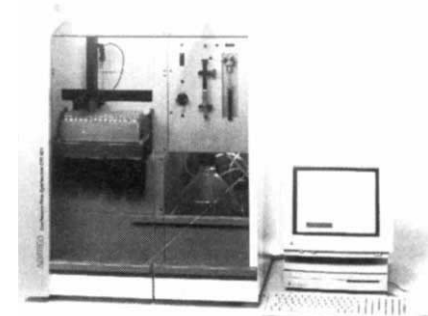
PCOS-1 from Large Scale Biology is a production-scale **centrifugal oligonucleotide synthesizer** that uses centrifugal force in combination with density differences between reagents to control flow through a short-path-length annular bed in a zonal centrifuge rotor (*Reader Service No. 103*).



PCOS-1: designed for the scale-up in the production of biopolymers.

The instrument is computer controlled and uses transparent rotary valves providing 47 reagent ports. Continuous multi-wavelength spectral and mass throughput data is provided during the run. And, as each of the reagents used (including the four bases) is distinguishable by spectral signature, LSB says that online data collected by PCOS-1 allows direct confirmation of the sequence amidites delivered, as well as the correct delivery of oxidizer, cap and deblock. Reaction kinetics, washing efficiency and sequence-position-specific effects can be studied quantitatively. X-Windows/Motif control software is included.

In need of high-purity long peptides? Abimed recommends its CFS 621 **continuous flow synthesizer**, which is designed to synthesize peptides with an optimal length



CFS 621: optimized for the synthesis of peptides of 20–80 amino acids in length.

of 20–80 amino acids (*Reader Service No. 104*). Active ultraviolet monitoring enables the continuous control of the whole synthesis and an automatic, PC-controlled optimization of synthesis parameters. Nevertheless, the user can intervene manually in the synthesis at any time. Control software allows users to work with standard synthesis protocols, or to define their own. For laboratories on a budget, Abimed offers the EPS 221 **economy peptide synthesizer**, which is capable of the unattended coupling of up to 45 amino acid derivatives distributed on 1–3 flow-through columns. With the EPS 221, the synthesis procedure has been optimized for a 100 μ mol scale regarding flow rates and solvent consumption.

Using the PSSM-8 **peptide synthesizer** from Shimadzu, up to eight peptides can be routinely synthesized in simultaneous mode (*Reader Service No. 105*). The instrument, which should be of use in epitope mapping studies and for studying structure-activity relationships, enables synthesis from 10 to

ADVERTISEMENTS

OLIGONUCLEOTIDES
 Synthesised, purified and shipped in 3 days
£2.50 per base
 (All inclusive)
0.2 µM scale (approx. 500 µg)
 - Antisense (completely or partially sulphurised)
 - Extensive range of labelling and modifications
 - Full design and analysis facilities
 - Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit
 King's College School of Medicine & Dentistry
 Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No.10

DNA SEQUENCING
 Using an ABI 373 Automated Sequencing System
£60/sample
 - PCR products - single stranded DNA
 - M13 inserts - double stranded DNA
 - plasmid inserts - solid phase sequencing
 - 300-400 bases of sequence

Please contact us for more information and a preparation protocol

Molecular Medicine Unit
 King's College School of Medicine & Dentistry
 Tel: 071 346 3126 Fax: 071-733 3877

Reader Service No.11

DNA, PEPTIDE SYNTHESIS
 PEPTIDE/PROTEIN SEQUENCING

Personalized Service
Guaranteed Quality
 Peptides: \$26.99/ residue
 Oligos: \$3.00 / base

TEL: 1-519-652-3758
FAX: 1-519-652-3816

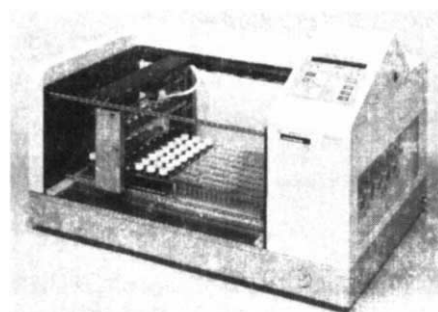
Reader Service No. 3

LIPOSOMES
 REPRODUCIBLE - HOMOGENEOUS -
 EXTREMELY RAPID

Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 1-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

BIACERM
 P.O. Box 211, 20, 0-7, 215 Munich

Reader Service No. 12



Eight peptides — one instrument. See the PSSM-8 solid-phase peptide synthesizer.

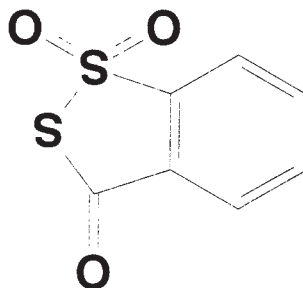
20 amino acid residue peptides in 1-1.5 days, says Shimadzu. It is possible to perform separation and characterization in less than 3 days. The fastest synthesis protocols are based on an average coupling time of 8-12 min. For every independent synthesis channel, from 5 to 50 µmol material can be synthesized; up to 400 µmol can therefore be synthesized in one run.

Kem-En-Tec's new MK3 automatic peptide synthesizer is optimized for Fmoc chemistry (Reader Service No. 106). The instrument can be preprogrammed for up to 50 amino acids. All cycle parameters can be defined by the operator. The company says that the instrument's design provides automatic dissolution of the amino acid and reduced solvent consumption.

Reagents

Synthetic oligonucleotides can be made resistant to nuclease attack by conversion to oligophosphorothioates (S-oligos) using an appropriate sulphurization reagent. If synthesis is performed with β-cyanoethyl phosphoramidite chemistry, the reagent of choice might be Beaucage reagent (3H-1,2-benzodithiole-3-one, 1,1-dioxide). Because

Beaucage reagent — BioGenex's sulphurization reagent for the fast synthesis of S-oligos.



it is readily soluble in acetonitrile, Bio-Genex Laboratories, which now supplies the reagent for the fast synthesis of S-oligos as a dried powder, says it can be used on automated synthesizers without the problem of clogging (Reader Service No. 107). The sulphurization reaction is said to take 30 s and is 96 per cent efficient in the conversion of phosphodiester linkages to phosphorothioates. Beaucage reagent is available in 1- and 2-g packets, ready for dissolution in anhydrous acetonitrile. Each order also includes a silanized bottle designed for use

with synthesizers from either Applied Biosystems or MilliGen.

Available now through Semat Technical — recombinant receptors from RBL (Reader Service No. 108). The new receptors, adrenergic, dopamine and muscarinic, can be used as reagents in drug discovery activities, including the screening of new molecular entities, and are available in a number of specific G protein-coupled receptor subtypes. They are supplied as 1 ml frozen cell membrane suspensions for 100 assay points.

Rec A protein for use in site-directed mutagenesis or targeted site-specific cleavage of DNA can now be obtained from Epicentre Technologies (Reader Service No. 109). A multi-functional DNA-binding protein isolated from *Escherichia coli*, Rec A is involved in both homologous recombination and post-replicative DNA repair mechanisms.

Labelling

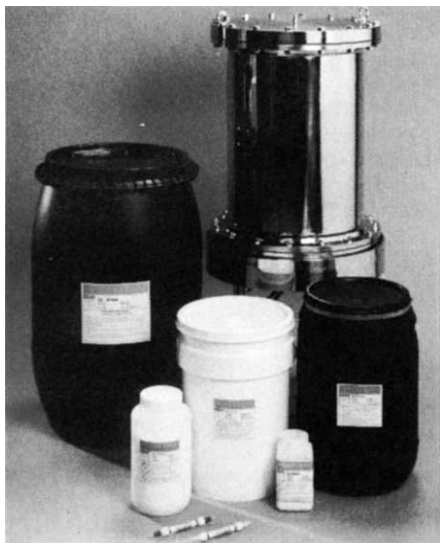
Amersham's oligonucleotide colour kit for use in nonradioactive *in situ* hybridization (ISH) is designed to provide low background and high signal-to-noise ratios (Reader Service No. 110). The kit introduces a tail of fluorescein-dUTP onto the 3'-end of an oligonucleotide. Labelling is said to take about an hour. And, says Amersham, there is no need to purify unincorporated nucleotide away from the labelled probe. Oligonucleotide probes for ISH can be used to study the regulation of gene expression, for the diagnosis of genetic disease, and for examining mRNA in complex heterogeneous tissue such as the central nervous system.

Perkin-Elmer's Applied Biosystems division has introduced three new fluorescent dye phosphoramidites, 6-FAM, HEX and TET, which are designed for the direct 5'-end labelling of primers and probes at coupling efficiencies greater than 90 per cent (Reader Service No. 111). Intended to facilitate gene analysis on the Model 373A DNA sequencing system (where they are compatible with the four-colour, one-lane detection technology) equipped with 672 Genescan software, these products can also be used for *in situ* hybridization, cell-uptake studies, target sequence identification, and other applications that would otherwise use radioactivity. Because the fluorescent label is added during oligonucleotide synthesis, Perkin-Elmer says that only an additional 20 min is required for the oligonucleotide purification cartridge to separate fluorescently labelled oligonucleotides from unlabelled species (conventional methods can often require more than 8 h of post-synthesis processing to add fluorescent labels and separate the product from both the unattached dye and unlabelled oligonucleotide). Purified yield from a 40-nmol synthesis is

1–2 ODU, said to be sufficient for approximately 650 PCR reactions.

■ Purification

Amberchrom is a family of high-capacity, **spherical polymeric resins** available from TosoHaas that can be used to purify a wide range of biopharmaceuticals (*Reader Service No. 112*). A new information hand-out is available from the company which characterizes and contrasts the performance of



Process resins for peptide purification.

Amberchrom resins for medium-pressure, preparative purification of different molecular weight polypeptides. Comparisons with traditional silica-based reversed-phase chromatography packings are included. Column longevity is illustrated by caustic cleaning studies.

Actigel ALD and Actigel Superflow are, according to the manufacturer, Sterogene Bioseparations, high-binding, fast coupling, **stable, activated resins** for proteins, antibodies (up to 50 mg ml⁻¹) and peptides (*Reader Service No. 113*). Used as an immunoaffinity support, the binding capacity is said to be two to three times higher than 'oriented' (hydrazide) or CNBr-coupling. With Actigel ALD, coupling can be performed at a pH most optimal for the protein, within pH 3–10 and at temperatures between 4 and 40 °C. Coupling takes about 2 h. Actigel Superflow is recommended for industrial applications with flow rates up to 3,000 cm h⁻¹. Actigel can be autoclaved or treated with NaOH.

Pharmacia Biotech now offers Sephasil C8 and Sephasil C18 in bulk and prepacked columns for high-performance, reversed-phase chromatography (RPC). These **silica-based matrices** are designed for high-resolution applications, such as purification of recombinant and synthetic peptides (up to 10–18 kilodaltons), peptide mapping and amino acid analysis (*Reader Service No. 114*). Sephasil C8 and C18 are silica-based,



Reversed-phase chromatography media in bulk and prepacked columns.

spherical beads with a nominal particle size of 5 µm and 12 µm, respectively. Prepacked Sephasil C8 5 µm and Sephasil C18 5 µm are available in SuperPac cartridge systems. Also available from the company — Source 15RPC and Resource RPC for reversed-phase chromatography of peptides, proteins and oligonucleotides. Providing an alternative to silica-based RPC matrices, Source media are characterized by a 15 µm monodispersed, polymeric matrix, where the uniformity of the beads is designed to give stable, packed beds and low back pressures. The operating pH range is 1–12. Resource RPC 1-ml columns (for screening applications) and 3-ml columns (for applications in which higher resolution is critical) are prepacked with Source 15RPC media.

Bio-Spin 6 and Bio-Spin 30 **chromatography columns** from Bio-Rad are prepacked, ready-to-use spin columns designed for the rapid desalting of biomolecules (*Reader Service No. 115*). In minutes, Bio-Rad says that the columns will clean up and purify probe, nick translation, plasmid and amplification reaction mixtures. They can also be



Bio-Rad's Bio-Spin columns

used to desalt and clean up labelled and unlabelled protein and peptide preparations, or to remove unbound dyes in dye-binding assays. These disposable (and autoclavable) columns are prepacked with Bio-Gel polyacrylamide gel filtration beads, a noncarbohydrate matrix. Bio-Spin 6 columns are said to have an exclusion limit of approximately 5 base pairs (bp), or 6,000 daltons for proteins; approximately 20 bp, or 40,000 daltons, for Bio-Spin 30 columns. The columns can be used with any 1.5-ml micro-test tubes or 12 × 75 mm test tubes in most tabletop clinical or low-speed floor-model centrifuges with swinging bucket rotors.

Gilson Medical Electronics has issued a new brochure that explains how its **HPLC system** automates the analysis, purification and reanalysis of synthetic peptides (*Reader*

Service No. 116). The system is completely automated, starting with the preparation of the crude peptide and ending with the collection of purified peptides or reanalysis of the purified material. In the Gilson peptide purification system, all modules are interfaced via the serial input/output channel. Using the 715 HPLC system controller software, the entire system can be controlled through a keyboard once the sampling injector and fraction collector have been programmed. To automate the system further, it can be configured with the 817 valve actuator for automatic column switching, a useful feature for switching between a preparative column for purification and an analytical column for reanalysis. Nine pump heads are available with flow rates ranging from 10 µl min⁻¹ to 200 ml min⁻¹. The brochure also includes a profile of the company's AMS 422 multiple peptide synthesizer. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more details, fill in the reader service card bound inside the journal. PCR is covered by patents owned by Hoffmann-La Roche.

ADVERTISEMENTS

Instead Synthesising Biomolecules by yourself just order them by fax:

TIB

MOLBIOL

Berlin +49 30-78 79 94 99
Genova +39 10-362 19 38

Reader Service No.23

PACEMAKER®

Custom peptide, conjugate and antibody production services

A range of quality services based on experience and expertise.

Peptide purities at >70%, >80%, or >95%; quantities from 1mg.

Hplc and mass spectrometric analysis for all peptides.

Conjugates prepared using carefully optimised procedures.

Protocols to comply with individual requirements.

Antibodies of defined specificity for utilisation in a variety of biochemical and immunological techniques.

Prices from £105 per peptide, £75 per conjugate.

For a brochure, further information or advice please contact the experts at:

AFFINITY Research Products Ltd

GPT Business Park, Technology Drive, Nottingham, NG9 2ND, United Kingdom

Tel: (+44/0)-602-436100; Fax: (+44/0)-602-436300

Reader Service No.12

PLEASE follow these guidelines so that your manuscript may be handled expeditiously.

Nature is an international journal covering all the sciences. Contributors should therefore bear in mind those readers who work in other fields and those for whom English is a second language, and write clearly and simply, avoiding unnecessary technical terminology. Space in the journal is limited, making competition for publication severe. Brevity is highly valued. One printed page of *Nature*, without display items, contains about 1,300 words.

Manuscripts are selected for publication according to editorial assessment of their suitability and reports from independent referees. They can be sent to London or Washington and should be addressed to the Editor. Manuscripts may be dealt with in either office, depending on the subject matter, and will where necessary be sent between offices by overnight courier. High priority cannot be given to pre-submission enquiries; in urgent cases they can be made in the form of a one-page fax. All manuscripts are acknowledged on receipt but fewer than half are sent for review. Those that are not reviewed are returned as rapidly as possible so that they may be submitted elsewhere without delay. Contributors may suggest reviewers; limited requests for the exclusion of specific reviewers are usually heeded. Manuscripts are usually sent to two or three reviewers, chosen for their expertise rather than their geographical location. Manuscripts accepted for publication are processed from the London office.

Nature requests authors to deposit sequence and crystallographic data in the databases that exist for this purpose.

Once a manuscript is accepted for publication, contributors will receive proofs in about 4 weeks. *Nature's* staff will edit manuscripts with a view to brevity and clarity, so contributors should check proofs carefully. Manuscripts are generally published 2–3 weeks after receipt of corrected proofs. *Nature* does not exact page charges. Contributors receive a reprint order form with their proofs; reprint orders are processed after the manuscript is published and payment received.

Categories of paper

Review Articles survey recent developments in a field. Most are commissioned but suggestions are welcome in the form of a one-page synopsis addressed to the Reviews Coordinator. Length is negotiable in advance.

Progress articles review particularly topical developments for a nonspecialist readership. They do not exceed 4 pages in length. Suggestions may be made to the Reviews Coordinator in the form of a brief synopsis.

Articles are research reports whose conclusions are of general interest and which are sufficiently rounded to be a substantial advance in understanding. They should not have more than 3,000 words of text (not including figure legends) or more than six display items and should not occupy more than five pages of *Nature*.

Articles start with a heading of 50–80 words written to advertise their content in general terms, to which editors will pay particular attention. The heading does not usually contain numbers, abbreviations or measurements. The introduction to the study is contained in the first two or three paragraphs of the article, which also briefly summarize its results and implications. Articles have fewer than 50 references and may contain a few short subheadings.

Letters to Nature are short reports of outstanding novel findings whose implications are general and important enough to be of interest to those outside the field. Letters should have 1,000 or fewer words of text and four or fewer display items. The first paragraph describes, in not more than 150 words and without the use of abbreviations, the background, rationale and chief conclusions of the study for the particular benefit of non-specialist readers. Letters do not have subheadings and contain fewer than 30 references.

Commentary articles deal with issues in, or arising from, research that are also of interest to readers outside research.

News and Views articles inform nonspecialist readers about new scientific advances, sometimes in the form of a conference report. Most are commissioned but proposals can be made in advance to the News and Views Editor.

Scientific Correspondence is for discussion of topical scientific matters, including those published in *Nature*, and for miscellaneous contributions. Priority is given to letters of fewer than 500 words.

Preparation of manuscripts

All manuscripts should be typed, double-spaced, on one side of the paper only. An original and four copies are required, each accompanied by artwork. If photographs are included, five sets of originals are required; for line drawings, one set of originals and four good-quality photocopies are acceptable. Reference lists, figure legends and tables should all be on separate sheets, all of which should be double-spaced and numbered. Three copies of relevant manuscripts in press or submitted for publication elsewhere should be included with submitted manuscripts, clearly marked as such. Five copies of revised and resubmitted manuscripts, labelled with their manuscript numbers, are required, together with five copies of a letter detailing the changes made.

Titles are brief and simple. Active verbs, numerical values, abbreviations and punctuation are to be avoided. Titles should contain one or two key words for indexing purposes.

Artwork should be marked individually and clearly with the author's name and, when known, the manuscript number. Ideally, no figure should be larger than 28 by 22 cm. Figures with several parts are to be avoided and are permitted only if the parts are closely related, either experimentally or logically. Unlettered originals of photographs should be provided. Suggestions for cover illustrations, with captions and labelled with the manuscript number, are welcome. Original artwork is returned when a manuscript cannot be published.

Protein/nucleotide sequences should ideally be in the three-letter and not the single-letter code for amino acids. One column width of *Nature* can accommodate 20 amino acids or 60 base pairs.

Colour artwork. A charge of £500 per page is made as a contribution towards the cost of reproducing colour figures. Inability to pay these costs will not prevent publication of essential colour figures if the circumstances are explained. Proofs of colour artwork may be sent to contributors under separate cover from their galley proofs.

Figure legends should not exceed 300 words and ideally should be shorter. The figure is described first, then, briefly, the method. Reference to a method published elsewhere is preferable to a full description. Methods are not described in the text.

References are numbered sequentially as they appear in the text, followed by those in tables and finally by those in figure legends. Only papers published or in the press are numbered and included in the reference list. All other forms of reference should be cited in the text as a personal communication, manuscript submitted or in preparation. Text is not included in reference lists. References are abbreviated according to the *World List of Scientific Periodicals* (Butterworths, London, 1963–65). The first and last page numbers are included; reference to books includes publisher, place and date.

Abbreviations, symbols, units and Greek letters should be identified the first time they are used. Acronyms should be avoided whenever possible and, if used, defined. Footnotes are not used in the text.

Acknowledgements are brief and appear after the reference list; grant and contribution numbers are not allowed.

Supplementary information is material relevant to Articles or Letters which cannot, for lack of space, be published in full, but which is available from *Nature* on request.

Submission. Manuscripts can be sent to the Editor at 4 Little Essex Street, London WC2R 3LF, UK or at 1234 National Press Building, Washington, DC 20045, USA. A telephone and fax number should be included. Manuscripts or proofs sent by air courier to London should be declared as 'manuscripts' and 'value \$5' to prevent the imposition of import duty and value-added tax.